Universidad Carlos III de Madrid
Escuela Politécnica Superior
Departamento de Tecnología Electrónica
Ingeniería Técnica de Telecomunicación:
Sistemas de Telecomunicación

Proyecto Fin de Carrera

DISEÑO DE UNA INTERFAZ GRÁFICA DE EVALUACIÓN DE ALGORITMOS DE FIRMA MANUSCRITA

Autor: Juan Fernández García-Obledo
Tutor: Oscar Miguel Hurtado
Julio 2010

“Who is it directed to?” said one of the jurymen.
“It isn’t directed at all,” said the White Rabbit; “in fact,
there’s nothing written on the outside.” he unfolded the paper as he spoke,
and added “It isn’t a letter, after all: it’s a set of verses.”
“Are they in the prisoner’s handwriting?” asked another of the jurymen.
“No, they’re not,” said the White Rabbit, “and that’s
the queerest thing about it.” (The jury all looked puzzled.)
“He must have imitated somebody else’s hand,” said
the King. (The jury all brightened up again.)
“Please your Majesty,” said the Knave, “I didn’t write it,
and they can’t prove I did: there’s no name signed at the end.”
“If you didn’t sign it,” said the King, “that only makes the matter worse.
You must have meant some mischief, or else you’d have
signed your name like an honest man.”

(Lewis Carroll, Alice in Wonderland)

ÍNDICE

LISTA DE FIGURAS

1

1.INTRODUCCIÓN

El estudio de sistemas biométricos está muy avanzado en la actualidad. Es muy importante la evolución de estos sistemas ya que pretenden identificar a un individuo a partir de un rasgo biológico propio como puede ser una huella dactilar o el iris, por ejemplo. Si lo pensamos un poco, cada persona es diferente a los demás y por tanto es lógico que se quiera buscar métodos automáticos de identificación con el estudio de los rasgos. Lo complicado será encontrar sistemas fiables antes de poderlos introducir en la sociedad.

La firma manuscrita es un rasgo de cada persona que tiene gran importancia en la sociedad. Con una firma concluimos una carta, asumimos un contrato, terminamos un cuadro,… en el fondo lo que queremos decir con nuestra firma es que el mensaje sobre lo que escribimos lo consideramos como nuestro. Esta firma por tanto merece ser analizada con detalle para poder verificar que una persona es realmente quien dice ser. La verificación de una firma implicará la aceptación del contenido de un documento u obra. Es importante conseguir entonces sistemas altamente eficientes sobre todo para temas relacionados con el comercio (firma de facturas) o el sector inmobiliario (firmas de hipotecas) por mencionar sólo unos pocos ya que la firma es algo que se utiliza casi a diario, es parte de nuestro D.N.I.

Existen algoritmos de análisis de firma manuscrita pero esta modalidad necesita todavía mucha investigación. Este proyecto surge con motivo del análisis de estos algoritmos. Se necesitan evaluar los errores para mejorar los algoritmos o para encontrar umbrales para poder decidir sobre la identidad de una persona.

Para comenzar este documento se tratará de avanzar brevemente las tareas que se han realizado en este proyecto.

Se mostrará el ámbito en el cual este proyecto se aplica y se relatarán los procedimientos y herramientas utilizadas para llevarlo a cabo y así alcanzar los objetivos que en este mismo capítulo también aparecerán. De esta manera se intentará poner al lector en situación para poder seguir la evolución del proyecto.

Al final de este capítulo introductorio se mostrará una estructuración de la memoria completa para facilitar el manejo de este documento.

1.1.MOTIVACIÓN

Este documento va a tratar el proyecto: **“Diseño de una interfaz gráfica de evaluación de algoritmos de firma manuscrita”**. En el departamento de Tecnología Electrónica de la Universidad Carlos III de Madrid, el Grupo Universitario de Tecnologías de la Identificación (G.U.T.I.) trabaja en diversos ámbitos de la biometría y uno de ellos es la identificación de personas por firma manuscrita. Se han desarrollado distintos algoritmos que permiten analizar distintas bases de datos de firmas de usuarios.

El presente trabajo va a utilizar estos algoritmos para presentar los resultados obtenidos del análisis de una base de datos. Se ha desarrollado una plataforma de trabajo gracias a la cual se pueden evaluar distintas bases de datos en base a los algoritmos de verificación de firma manuscrita. La aplicación permite simulaciones variando distintos parámetros para ver la evolución de un algoritmo. Pretende ser una herramienta de trabajo para los investigadores que puedan añadir mejoras a sus algoritmos o si es necesario se podrán añadir nuevos algoritmos. Esta aplicación mostrará gráficas de error en función de los distintos parámetros introducidos y de esta manera se podrá ver el rendimiento de un algoritmo o qué parámetros influyen más que otros.

La tutoría de este proyecto corresponde a D. Oscar Miguel Hurtado y la realización a D. Juan Fernández García-Obledo.

1.2.OBJETIVOS

El objetivo de este proyecto es desarrollar una aplicación que permita evaluar distintos algoritmos de firma manuscrita con distintas bases de datos de usuarios. Las firmas de estos usuarios estarán almacenadas en memoria siguiendo una estructura ya definida según la norma ISO/IEC 19794-7 Full Format.

La aplicación desarrollada va a permitir:

1. Introducir nuevas bases de datos de firmas de usuarios en memoria.
2. Eliminar bases de datos de firmas de usuarios.
3. Mostrar información de bases de datos de firmas de usuarios.
4. Mostrar información de los algoritmos de evaluación de firmas.
5. Seleccionar una base de datos de firmas de usuarios existente y un algoritmo para realizar una evaluación bajo una serie de parámetros.
6. Cargar resultados de simulaciones anteriores guardadas en ficheros con formato binario.
7. Guardar los resultados obtenidos de una simulación.

Para el desarrollo de esta aplicación se ha utilizado como herramienta el Borland Developer Studio 2006, la cual nos permite trabajar con C++ realizando Programación Orientada a Objetos (P.O.O.). Las funciones relativas a los algoritmos de los distintos algoritmos, así como las funciones de lectura de firmas las han desarrollado miembros del G.U.T.I.

1.3.ORGANIZACIÓN DEL DOCUMENTO

- **Capítulo 1:** Introducción y planteamiento de objetivos. Se intenta poner en situación al lector ante el propósito de este proyecto.

- **Capítulo 2:** Introducción a la Biometría. En este capítulo se da una breve historia de la Biometría con mención a los sistemas biométricos existentes, estándares referidos a la tecnología biométrica y aplicaciones.

- **Capítulo 3:** Firma manuscrita. El proyecto se centra en la firma manuscrita y por tanto este capítulo trata de analizar un poco su historia así como sus elementes o características. También se verá la importancia que tiene hoy en día y qué utilidad tiene en la sociedad.

- **Capítulo 4:** Algoritmos de firma manuscrita. Este capítulo presentará una descripción de los algoritmos utilizados en la evaluación de las firmas.

- **Capítulo 5:** Bases de datos de firma manuscrita. Relación y breve descripción de diversas bases de datos existente para el estudio.

- **Capítulo 6:** Normativa relativa a la firma manuscrita. Para el estudio de la firma manuscrita existen ciertas normas que por ejemplo definen la forma de almacenar una firma en una base de datos. En este capítulo se hará una breve descripción del formato de datos utilizado.

- **Capítulo 7:** Evaluación de bases de datos de firma manuscrita. Tipos de errores, representación gráfica y análisis de resultados. A la hora de analizar un algoritmo de verificación de firma será importante ver los distintos ratios de error a través de gráficas y se verá el significado de las mismas.

- **Capítulo 8:** Diseño, arquitectura y desarrollo. Descripción del desarrollo de las diversas funciones como lectura de una base de datos, cálculo de patrones de semejanza, cálculo de similitudes y simulaciones de un algoritmo. Análisis de los resultados obtenidos de cada función. Resultados obtenidos para algoritmos concretos y bases de datos concretas.

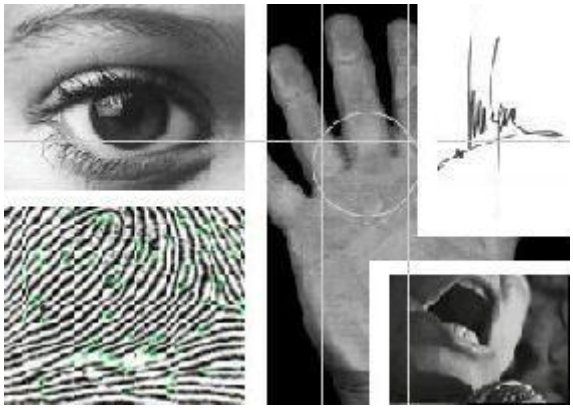
- **Capítulo 9:** Conclusiones y trabajos futuros. Una vez vistos los resultados con algoritmos y bases de datos actuales se analizan las conclusiones y se verá la línea de evolución de este campo de la biometría.

- **Capítulo 10:** Referencias

2

2. INTRODUCCIÓN A LA BIOMETRÍA

Cada individuo tiene características propias que le diferencian del resto. Somos únicos y esta característica nuestra se utiliza para mejorar nuestra vida para identificarnos. La Biometría es la ciencia que se dedica al estudio de los rasgos personales. Este capítulo trata de dar una visión general de esta ciencia que tiene muchas aplicaciones en la actualidad y muchos campos de investigación que están en desarrollo.



2.1. BIOMETRÍA

> **Biometría** (de bio- y –metría): estudio mensurativo o estadístico de los fenómenos o procesos biológicos.

Esta es la definición que el Diccionario de la Real Academia Española de la Lengua nos da de esta palabra.
Una definición más tecnológica y que nos servirá mejor es la siguiente: la **biometría** es una tecnología de identificación basada en el reconocimiento de una característica física e intransferible de las personas, como por ejemplo, la huella digital. La biometría es un excelente sistema de identificación de la persona que se aplica en muchos procesos debido a dos razones fundamentales, la seguridad y la comodidad.

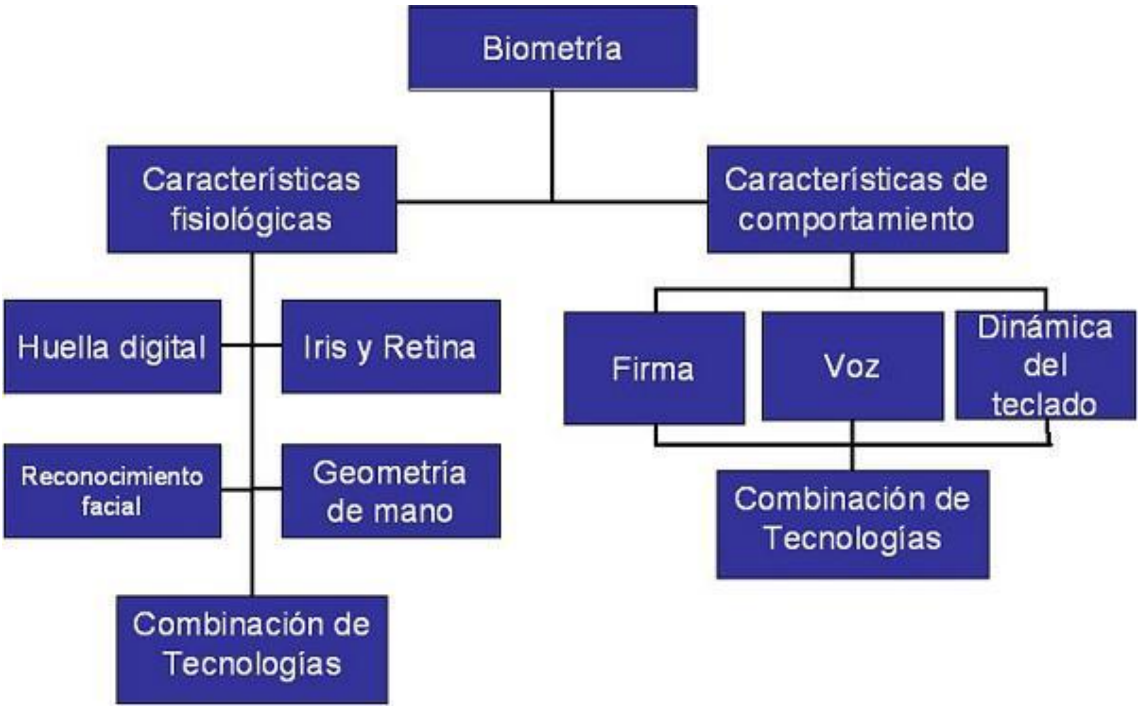


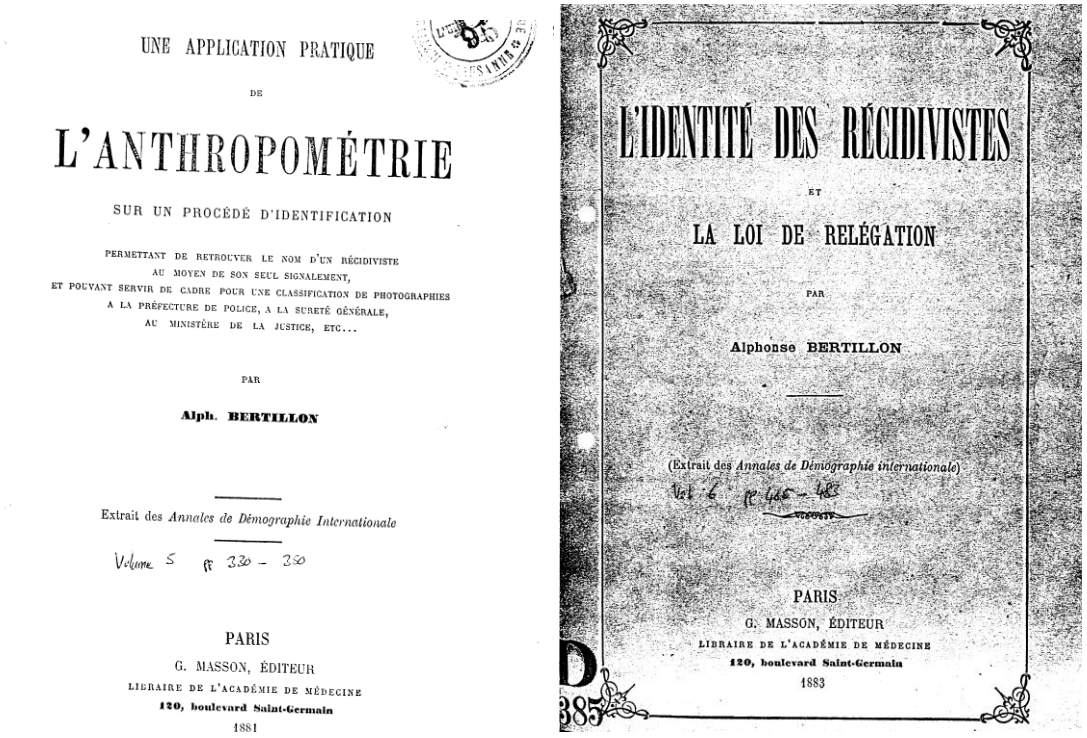
Figura 2.1: Características biométricas

2.1.1. HISTORIA

Hace ya bastantes años que se aplican técnicas de biometría para la identificación de las personas. Existen documentos que nos muestran cómo en siglos pasados han utilizado sistemas biométricos en contextos como la seguridad nacional entre los más importantes.

El explorador Joao de Barros (1496-1570) habló de que el primer ejemplo conocido de huellas dactilares se remonta a la época de China durante el siglo 14 dC. Los comerciantes chinos utilizan tinta para las huellas dactilares de los niños con fines de identificación.

Ya en el siglo XIX, con el crecimiento de las ciudades, se ve necesario el desarrollo de técnicas de identificación. Se quieren imponer castigos más severos a los presos reincidentes y para ello se analizan ciertos rasgos físicos de ellos para compararlos luego con una base de datos.


UNE APPLICATION PRATIQUE

DE

L'ANTHROPOMÉTRIE

SUR UN PROCÉDÉ D'IDENTIFICATION

PERMETTANT DE RETROUVER LE NOM D'UN RÉCIDIVISTE
AU MOYEN DE SON SEUL SIGNALEMENT,
ET POUVANT SERVIR DE CADRE POUR UNE CLASSIFICATION DE PHOTOGRAPHIES
A LA PRÉFECTURE DE POLICE, A LA SURETÉ GÉNÉRALE,
AU MINISTÈRE DE LA JUSTICE, ETC...

PAR

Alph. BERTILLON

Extrait des *Annales de Démographie Internationale*

Volume 5 p. 330 - 350

PARIS
G. MASSON, ÉDITEUR
LIBRAIRE DE L'ACADÉMIE DE MÉDECINE
120, boulevard Saint-Germain
1881

L'IDENTITÉ DES RÉCIDIVISTES
ET
LA LOI DE RELÉGATION

PAR

Alphonse BERTILLON

(Extrait des *Annales de Démographie internationale*)

Vol. 6 p. 465 - 485

PARIS
G. MASSON, ÉDITEUR
LIBRAIRE DE L'ACADÉMIE DE MÉDECINE
120, boulevard Saint-Germain
1883

Figura 2.2.: Libros editados en el siglo XIX sobre la biometría

En el siglo XX, gracias al desarrollo de los ordenadores ha surgido nuevos sistemas biométricos más modernos y especializados. Existen sistemas electrónicos que permiten la lectura de huellas digitales, o tabletas que pasan a un ordenador una firma manuscrita.

El siguiente cuadro trata de resumir un poco la evolución temporal de la biometría sobre todo a partir de finales del siglo XIX:

1890	Alphonse Bertillon al servicio de policía de París estudió la mecánica del cuerpo y las medidas para ayudar a identificar a los delincuentes. La policía utilizó su método, el método Bertillonage, hasta que se identificó falsamente algunos temas.
Principios del siglo XX	Karl Pearson, matemático, estudió la investigación biométrica en el University College de Londres. Hizo importantes descubrimientos en el campo de la biometría a través de estudio de la historia y la correlación estadística, que aplicó a la evolución de los animales. Su obra histórica incluye el método de momentos, el sistema de las curvas de Pearson, la correlación y la prueba de chi-cuadrado
Década de 1960 y 1970	Los procedimientos de la firma autentificación biométrica se han desarrollado, pero el campo se mantuvo fijo hasta biométricos de los organismos militares y de seguridad investigado y desarrollado más allá de la tecnología biométrica de huellas digitales.
11-09-2001	Los pasaportes biométricos forman parte de una exigencia de seguridad establecida por el Department of Homeland Security (DHS) tras los ataques terroristas del 11 de septiembre de 2001. Entre otros datos, deben incluir la impresión de la huella dactilar, foto del iris del portador y un microchip insertado en una de sus páginas donde se almacenará la información del rostro de la persona que porta el documento.
2004	Telefónica I+D eligió la celebración del congreso IST 2004 para presentar oficialmente los resultados de BioSec, su iniciativa para investigar las técnicas biométricas, que ofrecen mecanismos digitales para conocer a las personas y verificar su identidad a partir de sus características fisiológicas (imágenes del iris o del rostro, voz, huella dactilar o geometría de la mano, entre otras). El objetivo de Biosec es mejorar las características de seguridad de los sistemas de identificación biométricos. En el congreso, Telefónica I+D ha mostrado las ventajas de integrar la biometría para acceder a redes de datos, como alternativa al uso de claves, ya que es imposible duplicar las características biométricas.
2006	Identificación del iris en lugar de número PIN para acceder al móvil. Este sistema ya está disponible para teléfonos móviles. La seguridad y la comodidad son los factores que, según un estudio de LogicaCMG, han llevado a los europeos a una mayoritaria aceptación de la biometría.
Abril 2010	La investigadora española Celia Sánchez Ramos obtiene el Gran Premio del Jurado del trigésimo octavo Salón Internacional de las Invenciones Técnicas y Nuevos Productos de Ginebra por un sistema de identificación de personas basado en la Biometría Ocular. La invención permite comparar en un segundo la topografía interna de la córnea con la que está registrada en un banco de datos, sin causar efectos secundarios.

2.1.2. SISTEMAS BIOMÉTRICOS

¿Qué hace que un rasgo sea considerado biométrico? Ante todo, es necesario que esté presente en cualquier persona susceptible de ser identificada (**universalidad**). Además, la característica a medir no debe cambiar sustancialmente con el paso del tiempo (**permanencia**). Por otra parte, resulta crucial que la probabilidad de que los patrones de dos sujetos coincidan sea muy baja (**unicidad**). Sin olvidar que el rasgo debe ser **cuantificable**, **accesible** y, en la medida de lo posible, fácil de obtener.

Entenderemos por *sistema biométrico* a un sistema automatizado que realiza labores de biometría. Es decir, un sistema que fundamenta sus decisiones de reconocimiento mediante una característica personal que puede ser reconocida o verificada de manera automatizada.

En base a la obtención de un sistema biométrico con utilidad práctica podremos hablar de ciertas características que se considerarán a la hora de su diseño:

1. El *desempeño*, que se refiere a la exactitud, la rapidez y la robustez alcanzada en la identificación, además de los recursos invertidos y el efecto de factores ambientales y/u operacionales. El objetivo de esta restricción es comprobar si el sistema posee una exactitud y rapidez aceptable con un requerimiento de recursos razonable.
2. La *aceptabilidad*, que indica el grado en que la gente está dispuesta a aceptar un sistema biométrico en su vida diaria. Es claro que el sistema no debe representar peligro alguno para los usuarios y debe inspirar “confianza” a los mismos. Factores psicológicos pueden afectar esta última característica. Por ejemplo, el reconocimiento de una retina, que requiere un contacto cercano de la persona con el dispositivo de reconocimiento, puede desconcertar a ciertos individuos debido al hecho de tener su ojo sin protección a un “aparato”. Sin embargo, las características anteriores están subordinadas a la aplicación específica. En efecto, para algunas aplicaciones el efecto psicológico de utilizar un sistema basado en el reconocimiento de características oculares será positivo, debido a que este método es eficaz implicando mayor seguridad.
3. La *fiabilidad*, que refleja cuán difícil es burlar al sistema. El sistema biométrico debe reconocer características de una persona viva, pues es posible crear dedos de látex, grabaciones digitales de voz, prótesis de ojos, etc. Algunos sistemas incorporan métodos para determinar si la característica bajo estudio corresponde o no a la de una persona viva. Los métodos empleados son ingeniosos y usualmente más simples de lo que uno podría imaginar. Por ejemplo, un sistema basado en el reconocimiento de iris revisa los patrones característicos en las manchas de éste, un sistema infrarrojo para chequear las venas de la mano detecta flujos de sangre caliente y lectores ultrasonido para huellas dactilares revisan estructuras subcutáneas de los dedos.

Los sistemas biométricos poseen tres componentes básicos. El primero se encarga de la adquisición analógica o digital de algún indicador biométrico de una persona, como por ejemplo, la adquisición de la imagen de una huella dactilar mediante un escáner. El

segundo maneja la compresión, procesamiento, almacenamiento y comparación de los datos adquiridos con los datos almacenados. El tercer componente establece una interfaz con aplicaciones ubicadas en el mismo u otro sistema.

La arquitectura típica de un sistema biométrico se presenta en la siguiente figura:

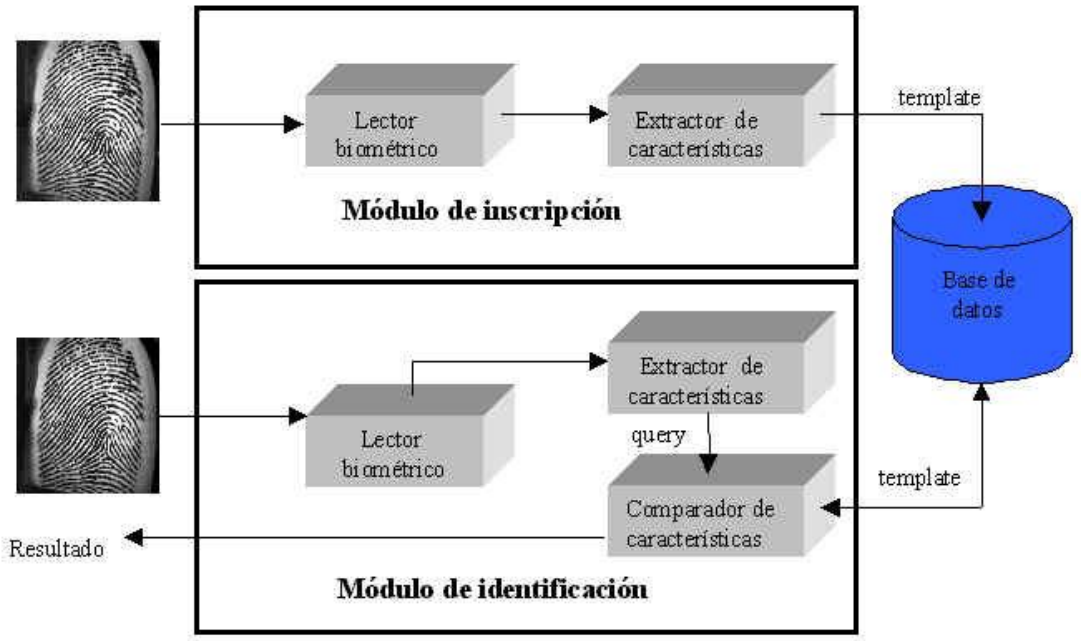


Figura 2.3: Arquitectura de un sistema biométrico para identificación personal

En la figura se distinguen dos módulos diferentes:

1. El *módulo de inscripción* se encarga de adquirir y almacenar la información proveniente del indicador biométrico con el objeto de poder contrastar a ésta con la proporcionada en ingresos posteriores al sistema. Las labores ejecutadas por el módulo de inscripción son posibles gracias a la acción del lector biométrico y del extractor de características.
2. El *módulo de identificación* extrae características representativas del indicador para luego compararlas con la información almacenada en una base de datos que previamente ha sido completada por el módulo de inscripción.

Un sistema biométrico puede operar en dos modos:

1. Modo de verificación: comprueba la identidad de algún individuo comparando la característica sólo con registros guardados de ese individuo. El modo de verificación responderá a la pregunta: ¿eres tú quien dices ser?
2. Modo de identificación: descubre a un individuo mediante una comparación exhaustiva con todos los registros guardados de distintos individuos. El modo de identificación responderá a la pregunta: ¿quién eres tú?

2.1.3. ESTÁNDARES ASOCIADOS A TECNOLOGÍAS BIOMÉTRICAS

En los últimos años se ha notado una preocupación creciente por las organizaciones regulatorias respecto a elaborar estándares relativos al uso de técnicas biométricas en el ambiente informático. Esta preocupación es reflejo del creciente interés industrial por este ámbito tecnológico, y a los múltiples beneficios que su uso aporta. No obstante, aún la estandarización continúa siendo deficiente y como resultado de ello, los proveedores de soluciones biométricas continúan suministrando interfaces de software propietarios para sus productos, lo que dificulta a las empresas el cambio de producto o vendedor.

A nivel mundial el principal organismo que coordina las actividades de estandarización biométrica es el Sub-Comité 37 (SC37) del Joint Technical Committee on Information Technology (ISO/IEC JTC1), del International Organization for Standardization (ISO) y el International Electrotechnical Commission (IEC).

En Estados Unidos desempeñan un papel similar el Comité Técnico M1 del INCITS (International Committee for Information Technology Standards), el National Institute of Standards and Technology (NIST) y el American National Standards Institute (ANSI).

Existen además otros organismos no gubernamentales impulsando iniciativas en materias biométricas como: Biometrics Consortium, International Biometrics Groups y BioAPI. Este último se estableció en Estados Unidos en 1998 compuesto por las empresas Bioscrypt, Compaq, Iridiam, Infineon, NIST, Saflink y Unisis. El Consorcio BioAPI desarrolló conjuntamente con otros consorcios y asociaciones, un estándar que promoviera la conexión entre los dispositivos biométricos y los diferentes tipos de programas de aplicación, además de promover el crecimiento de los mercados biométricos.

Algunos de los estándares más importantes son:

- Estándar ANSI X.9.84: creado en 2001 por la ANSI (American National Standards Institute) y actualizado en 2003, define las condiciones de los sistemas biométricos para la industria de servicios financieros haciendo referencia a la transmisión y almacenamiento seguro de información biométrica y a la seguridad del hardware asociado.
- Estándar ANSI/INCITS 358: creado en 2002 por ANSI y BioAPI Consortium, presenta una interfaz de programación de aplicación que garantiza que los productos y sistemas que cumplen este estándar son interoperables entre sí.
- Estándar NISTIR 6529: también conocido como CBEFF (Common Biometric Exchange File Format), es un estándar creado en 1999 por NIST y Biometrics Consortium que propone un formato estandarizado (estructura lógica de archivos de datos) para el intercambio de información biométrica.
- Estándar ANSI 378: creado en 2004 por la ANSI, establece criterios para representar e intercambiar la información de las huellas dactilares a través del uso de minucias. El propósito de esta norma es que un sistema biométrico

dactilar pueda realizar procesos de verificación de identidad e identificación, empleando información biométrica proveniente de otros sistemas.

- Estándar ISO 19794-2: creado en 2005 por la ISO/IEC con propósitos similares a la norma ANSI 378, respecto a la que guarda mucha similitud.
- Estándar PIV-071006: creado en 2006 por el NIST y el FBI en el contexto de la norma FIPS 201 del gobierno de Estados Unidos, establece los criterios de calidad de imagen que deben cumplir los lectores de huellas dactilares para poder ser usados en procesos de verificación de identidad en agencias federales.

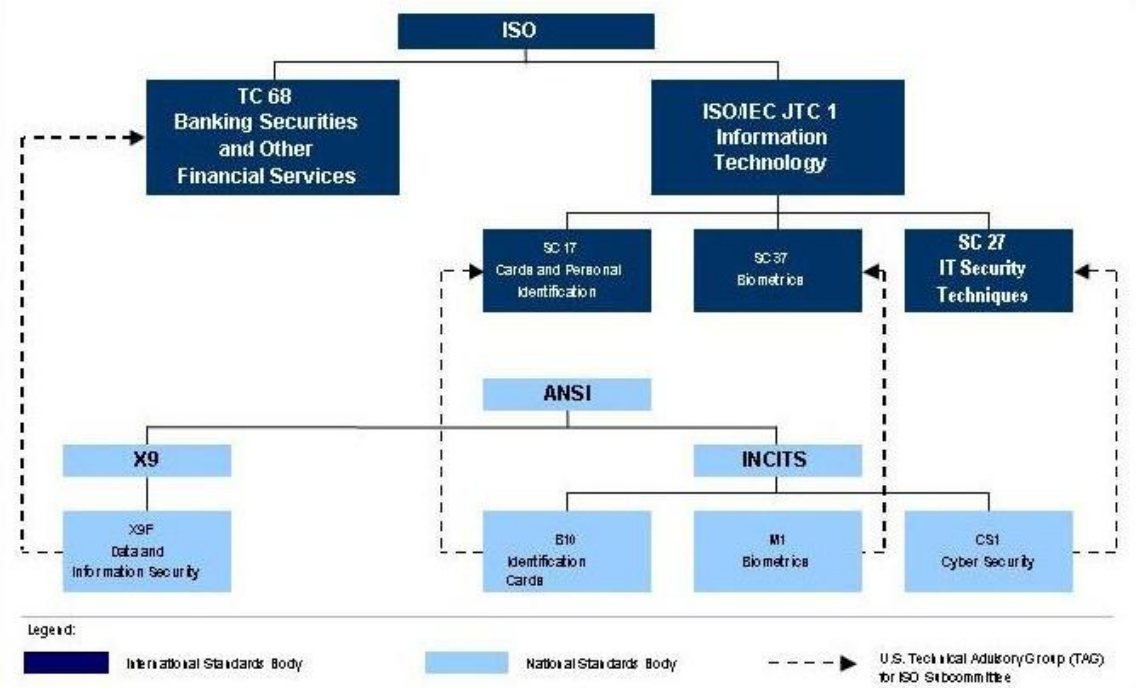


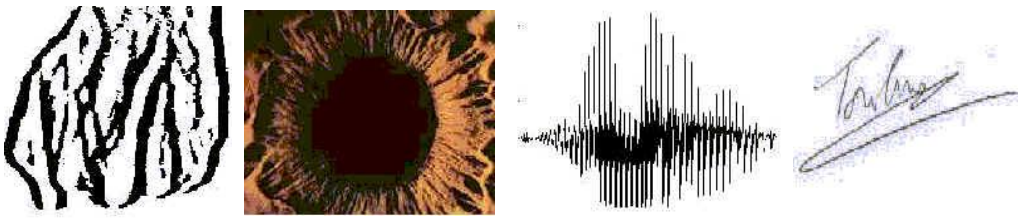
Figura 2.4.: Esquema organizativo de los comités

2.1.4. APLICACIONES

En la actualidad existen sistemas biométricos que basan su acción en el reconocimiento de diversas características. Algunas de las técnicas biométricas más conocidas están basadas en los siguientes indicadores biométricos:

1. Rostro,
2. Huellas dactilares,
3. Geometría de la mano,
4. Venas de las manos,
5. Iris,
6. Voz,
7. Firma.

(a) (b) (c)



(d) (e) (f) (g)

Figura 2.5.: Técnicas biométricas actuales: (a) Rostro, (b) Huella dactilar, (c) Geometría de la mano, (d) Venas de la mano, (e) Iris, (f) Voz e (g) Firma

En la actualidad existen fabricantes que suministran sus productos para su uso en distintas facetas de la vida. Existen también grupos de investigación que desarrollan nuevos algoritmos y productos para la identificación biométrica.

Entre los fabricantes destacamos:

Kimaldi : “Fabricamos y distribuimos sistemas de Control de Acceso y Control de Presencia, lectores biométricos de huella digital y biométricos vasculares,

RFID, lectores e impresoras de tarjetas."(Texto sacado de la página de la empresa www.kimaldi.com).

LG Group: LG es una empresa de Corea del Sur que fabrica productos electrónicos, teléfonos móviles y productos petroquímicos y opera con filiales como LG Electronics. En algunos de sus productos se han introducido sistemas biométricos como lectores de huellas digitales.

Siemens: "El Centro Biométrico de Siemens IT Solutions and Services se configura como especialista en soluciones de altos requerimientos y capaz de recurrir a unos conocimientos especializados completos. Ofrecemos una amplia gama de productos y soluciones comprobadas y verificadas en el mercado, tales como sistemas de control de fronteras y pasaportes para la seguridad pública o sistemas para la seguridad corporativa y el acceso a edificios y equipos de TI protegidos biométricamente."(Texto sacado de la página de la empresa www.swe.siemens.com).

L-1 Identity Solutions: "With the trust and confidence in individual identities provided by L-1, governments and businesses around the world are better protecting the public against terrorism, crime and theft fostered by fraudulent ID."(Texto sacado de la página de la empresa www.l1id.com).

Existen varios grupos de investigación biométrica en España, entre ellos destacamos:

GAVAB: El Grupo de Algorítmica para la Visión Artificial y la Biometría se fundó en 2006 en la Universidad Rey Juan Carlos. Se fundó originalmente para la investigación de técnicas de Visión Artificial y Biometría. Actualmente han ampliado el campo estudiando temas como el procesamiento del lenguaje

natural o el procesamiento de altas prestaciones (imágenes 3D, vídeo en tiempo real,…).

GUTI: El Grupo Universitario para las Tecnologías de Identificación fue fundado inicialmente por el Dr. José L.Zoreda en el año 1989 en la ETSIT de la Universidad Politécnica de Madrid. En el año 2000 se pasó a la Universidad Carlos III de Madrid dirigida por el Dr. Raúl Sánchez-Reillo. Las líneas de investigación del GUTI están dirigidas a los dispositivos de identificación (centrándose en tarjetas inteligentes y RFID), la seguridad de la información y la biometría.

3

3. FIRMA MANUSCRITA

> **Firma:** nombre y apellido, o título, que una persona escribe de su propia mano en un documento, para darle autenticidad o para expresar que aprueba su contenido. (Definición Diccionario RAE).

La principal razón por la que se considera interesante trabajar sobre el método de autenticación basado en firma manuscrita es su gran aceptación social, dado que la gente lo ha venido usando durante muchos años. Se trata de conseguir que la verificación manual, costosa y propensa a errores, sea sustituida por una verificación automática más fiable y de bajo coste.

Se pueden distinguir a grandes rasgos dos tipos de técnicas para la verificación de firmas según el método utilizado para la obtención de la misma. Estos métodos serían:

- Métodos estáticos, que no tienen en cuenta la velocidad con la que se traza la firma, tan sólo la imagen resultante del acto de firmar (rúbrica).
- Métodos dinámicos, que sí tienen en cuenta las velocidades con las que se realizan los trazos, analizando las señales temporales obtenidas por los dispositivos de captura (posición x e y, presión, inclinaciones).

Si se considera a la firma como un conjunto de signos, se puede distinguir una doble función de la misma: la primera es una función *identificadora* de la persona, puesto que determina su personalidad; y luego también está la función de *autenticación* que consiste en el proceso por medio del cual se revelan algunos aspectos de la identidad de una persona. El que firma además de expresar su consentimiento tomará como suyo el mensaje.

3.1. HISTORIA

En la Antigua Roma había una ceremonia llamada *Manufirmatio*, que consistía en que una vez que se lee un documento, se pone extendido sobre la mesa del escribano y una vez que el lector pasa la mano abierta sobre él, pone el nombre, signo o una o tres cruces, haciéndolo después los testigos. Esto no era un requisito para cada lectura de documentos, más bien era parte del espectáculo para culminar la lectura de algunos documentos. Esta ceremonia simbolizaba autenticidad y compromiso.

En la Edad Media a los documentos se les ponía una cruz añadiendo algunas letras y signos. Esto era utilizado como firma. Debido a la gran cantidad de gente que no sabía leer ni escribir, se reemplaza esta firma por el uso de sellos.

Figura 3.1.: Documento sellado

La diferenciación entre “firmas” y “signos” hizo que se empezase a entender que aquéllas eran, más que simples “signos”, la inscripción manuscrita del nombre o de los apellidos. Con el transcurso del tiempo la firma se fue consagrando como un símbolo de identificación y de enlace entre el autor de lo escrito o estampado y su persona.

3.2. CARACTERÍSTICAS

Cuando un autor ha culminado su obra, para poner de manifiesto que le pertenece, pone una “firma” o un “signo” que puede ser su nombre (real o artístico) o un dibujo, pero será algo que le sirva para diferenciarse del resto de las personas.
También puede suceder que alguien deba dar permiso ante una circunstancia o deba asumir las clausulas de un contrato. Para ello también se firman documentos y de esta manera el autor de la firma asume el contenido de un mensaje y lo toma como suyo para darle cumplimiento. Esto lo vemos en la firma de documentos administrativos, permisos, contratos, etc.
En algunos casos se exige la presencia de un perito para probar que la persona que firma es quien dice que es, ya que es posible que aparezca un documento firmado por alguien que ha falsificado la firma.

Por tanto se pueden distinguir las siguientes características que definen una firma manuscrita:

- Identificativa: con ella se identifica al autor de un documento u obra.
- Declarativa: el autor asume el contenido del mensaje. En la conclusión de un contrato la firma implica la voluntad de obligarse a cumplirlo.
- Probatoria: sirve para identificar si efectivamente ha sido esa persona la que ha realizado la firma.

3.3. ELEMENTOS

Existen distintos elementos que se tendrán en cuenta en el estudio de una firma manuscrita:

- Elementos formales: son los elementos materiales que están en relación con los procedimientos utilizados para firmar. La firma es un signo personal ya que debe ser hecha por el mismo firmante de su propia mano.
- El animus sigmandi: es el elemento intencional o intelectual de la firma. Es la voluntad de asumir lo que se firma.
- Elementos funcionales: en la introducción de este capítulo ya se comentó algo sobre las funciones de la firma manuscrita:
    - Función identificadora: expresa la identidad, aceptación y autoría del firmante. No es método de autenticación 100% fiable ya que el documento podría ser modificado después de una firma.
    - Función de autenticación: el autor de la firma da su consentimiento y hace suyo el mensaje.

3.4. IMPORTANCIA

La firma ha sido usada para validar la autenticidad de transacciones bancarias u otros documentos legales desde hace muchos años. En esta época, donde se está evolucionando mucho en las Tecnologías de la Información, los papeles se están sustituyendo por documentos electrónicos. A pesar de ello, en la actualidad todavía se tiene que firmar a mano ya que no se ha encontrado otra manera efectiva de firmar los documentos electrónicos y garantizar su integridad.

La firma forma parte de la identidad de la persona. En nuestro carné de identidad tenemos una imagen de nuestra firma que en algunos sitios es pedida para comprobar a simple vista que somos nosotros. Sería interesante que nuestra firma sea almacenada en nuestro carné también digitalmente para luego poder hacer una comprobación automática de la misma.

El estudio de la firma manuscrita trata de hacer la vida de las personas más cómoda y sobre todo más segura. Cualquier persona puede ver nuestra firma y entrenarla para suplantar nuestra identidad. Pero hay algunos aspectos que son inimitables o por lo menos muy difíciles de copiar como puede ser la velocidad de firma. Podrán copiar nuestra forma pero no se hará a la misma velocidad. Este será uno de los aspectos que se analizarán para la identificación de personas a través de firma manuscrita.

3.5. APLICACIONES

En el apartado anterior se ha hablado sobre el uso de la firma en transacciones bancarias o en otros documentos legales, como puede ser un contrato laboral, un certificado médico, una hipoteca,…

La firma es aplicada entonces en muchos aspectos de nuestra vida. Cuando se paga con tarjeta de crédito se pide firmar la factura o cuando se va al banco a pedir un crédito se solicita poner nuestra firma para dar conformidad a lo que se expone.

La firma también es usada en el mundo del arte. En las pinturas se puede observar firmas de sus autores que ponen su nombre real o artístico para que el público pueda reconocer su obra.

Figura 3.2.: Detalle del cuadro “El Naufrago” de Asensio Juliá

Como conclusión de este apartado y del anterior que hablaba sobre la importancia de la firma, podemos ver que realmente hay un interés en el estudio de la firma manuscrita por parte de los distintos sectores de la industria. Todavía no se ha implantado totalmente la documentación electrónica y se necesita firmar manualmente muchos documentos. Esta firma lo que viene a significar es la identidad de la persona y su consentimiento sobre lo que contiene el documento.

Hoy en día el reconocimiento de la firma manuscrita es una realidad que está muy introducida en nuestra vida. Empresas como SOFTPRO[1] ofrece tecnología biométrica para evitar el fraude utilizando técnicas de reconocimiento estático y dinámico para sacar las características de una firma y esta tecnología es aplicada en transacciones económicas por ejemplo (Fraud Prevenition Solutions). También KECRYPT[2] que también lucha por la seguridad en la toma de firmas en documentos para ello ofrece productos como KeSignature para identificar a los usuarios con reconocimiento dinámico de la firma.

[1] http://www.signplus.com/en/

[2] http://www.kecrypt.com/index.php

3.6. BIOMETRÍA EN FIRMA MANUSCRITA

En el capítulo 2 de este documento se habla acerca de la Biometría. Uno de los campos de investigación de la misma se centra en la firma manuscrita.

Existe en el mercado ya algunos sistemas biométricos basados en firma manuscrita. Aún así, todavía se está investigando más en mejorar algoritmos o incluir nuevos algoritmos.

Se necesitará de una base de datos para el entrenamiento de estos sistemas y para ello existen aparatos como las tabletas electrónicas que permiten digitalizar una firma manuscrita.

Figura 3.3.: tableta digitalizadora

Dependiendo del modo en que se extrae la firma, se pueden distinguir dos tipos:

- **Firma Estática**: corresponde a la digitalización de una firma a partir de una copia obtenida en papel. Se la conoce también como firma *off-line*.
- **Firma Dinámica**: corresponde a las que se obtienen a través de dispositivos que permitan sacar características como velocidad o presión del lápiz. Se la conoce también como firma *on-line*.

A la hora de analizar un sistema biométrico basado en firma manuscrita debemos diferenciar dos funciones:

- **Identificación**: es el proceso en el cual se ve quién es el individuo comparando la firma con las muestras almacenadas en la base de datos.
- **Verificación**: es el proceso que sirve para comprobar que un individuo es realmente quien dice ser que es. Se extrae de la base de datos la firma y se comparan.

Para la evaluación de este sistema biométrico se necesitará comparar con firmas genuinas aportadas por el individuo y firmas falsas de posibles imitadores. De esta

manera se verá el porcentaje de aciertos respecto al total de las identificaciones realizadas para ver su eficiencia. Por tanto se pueden distinguir hasta tres tipos de usuarios de un sistema biométrico:

- El *cliente* o usuario que aporta firmas genuinas.
- El *impostor casual* que intenta suplantar la identidad del cliente pero que no conoce los rasgos de la firma.
- El *imitador* que sí ha tenido acceso a la observación de la firma del cliente y pretende suplantar su identidad.

A través de distintas bases de datos de firmas manuscritas se podrá ver la evolución de los distintos sistemas biométricos para poder mejorarlos. En este proyecto se observan los resultados obtenidos de dos algoritmos de verificación de firma manuscrita como son el DTW y el GMM que serán analizados en el siguiente capítulo y con unas bases de datos conocidas como SVC2004, MCyT y MyIDea que son analizadas en el capítulo 5 de este documento[3].

[3] J. Ortega-Garcia, et al., "MCYT baseline corpus: a bimodal biometric database," Vision, Image and Signal Processing, IEE Proceedings -, vol. 150, pp. 395-401, 2003.
D. Y. Yeung, et al., "SVC2004: First international signature verification competition," Biometric Authentication, Proceedings, vol. 3072, pp. 16-22, 2004.
B. Dumas, et al., "MyIdea - Multimodal Biometrics Database, Description of Acquisition Protocols," in Third COST 275 Workshop (COST 275), Hatfield (UK), 2005, pp. 59-62.
O. Miguel-Hurtado, et al., "A new algorithm for signature verification system based on DTW and GMM," in Security Technology, 2008. ICCST 2008. 42nd Annual IEEE International Carnahan Conference on, 2008, pp. 206-213.
O. Miguel-Hurtado, et al., "On-Line Signature Verification by Dynamic Time Warping and Gaussian Mixture Models," in Security Technology, 2007 41st Annual IEEE International Carnahan Conference on, 2007, pp. 23-29.

4

4. ALGORITMOS DE FIRMA MANUSCRITA

Una firma manuscrita puede ser digitalizada tomando algunas características de la misma. Existen elementos electrónicos que permiten grabar las firmas. La información que se guarda consiste en datos numéricos que pueden ser tratados a través de funciones. Para ello se desarrollan algoritmos realizan transformaciones con los datos para llegar al resultado deseado de identificación. En este capítulo se presentan los dos algoritmos que se han utilizado para la realización del proyecto.

4.1. DTW (DYNAMIC TIME WARPING)[4]

El algoritmo de Alineamiento Dinámico Temporal (DTW) es muy utilizado en el reconocimiento del habla. Tiene por objeto buscar un eje temporal común para permitir la comparación de dos señales.
El algoritmo trata de realizar un alineamiento temporal entre las dos señales a comparar. Uno de los ejes corresponde al patrón (p,pattern) y el otro es la muestra (s, sample). Se crea una matriz de distancias entre cada par de puntos del patrón y la muestra y se busca sobre esa matriz un camino óptimo.

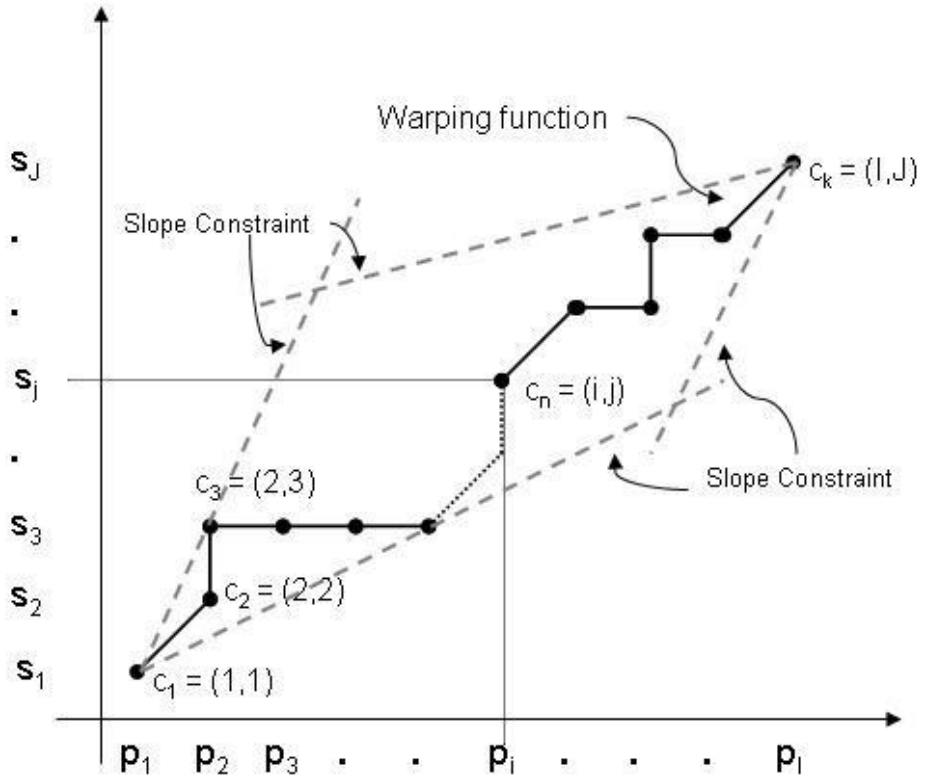


Figura 4.1.: Búsqueda del camino óptimo

La información que se obtiene de una firma mediante una tableta gráfica será la posición x, la posición y, la presión, el azimuth y la inclinación.

[4] O. Miguel-Hurtado, et al., "A new algorithm for signature verification system based on DTW and GMM," in Security Technology, 2008. ICCST 2008. 42nd Annual IEEE International Carnahan Conference on, 2008, pp. 206-213.

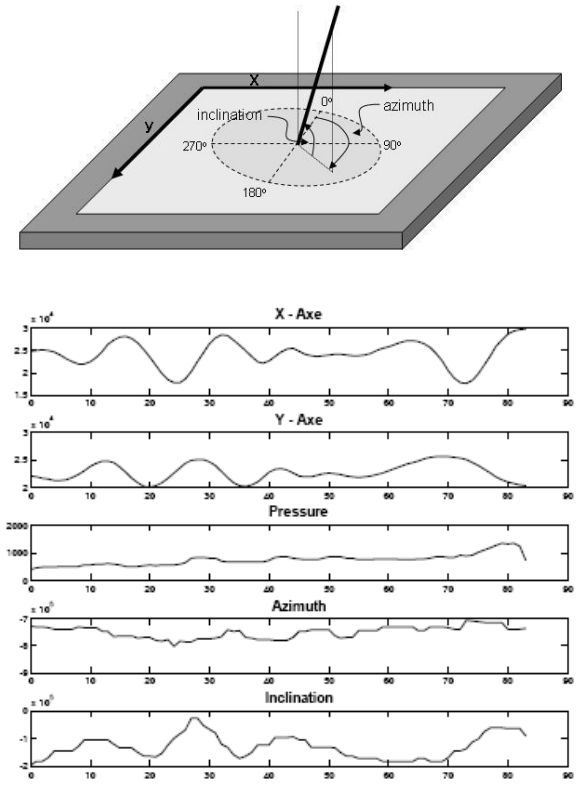


Figura 4.2.: Esquema de la información obtenida de la tableta digital

Antes de calcular el camino óptimo se hace un pre-procesado a la señal que consiste en los siguientes puntos:

1. Filtrado.
2. Equi-espaciado por Interpolación Lineal.
3. Normalización temporal.
4. Normalización espacial.
5. Normalización en tamaño.

Una vez que se han realizado los 5 puntos anteriores sobre la señal adquirida de la tableta digital, se procede a alinear temporalmente el patrón con la muestra. A través de la combinación de las señales X e Y se encuentra el camino optimo que minimiza la distancia entre el patrón y la firma. Utilizando este camino, se realiza la alineación de las 5 señales originales (X,Y,Az,El,P).

Cuando se hayan alineado temporalmente las señales originales, se procede a calcular las pseudo-distancias y se calculan como la media de la diferencia absoluta entre las señales (originales y derivadas) del patrón y la muestra a través de la siguiente fórmula:

$$ps_s = \frac{\sum_{i=1}^{N} \left| s_{input}(i) - s_{pattern}(i) \right|}{N}$$

El valor de N viene fijado del proceso de interpolación previo.

Para el cálculo de señales derivadas se utiliza la siguiente fórmula de regresión:

$$reg(z(t), O) = \frac{\sum_{k=1}^{O} k(z(t+k) - z(t-k))}{2\sum_{k=1}^{O} k^2}$$

Para el cálculo de las pseudo-distancias se utilizan las siguientes 25 señales:

1. Coordenada $x(t)$
2. Coordenada $y(t)$
3. Velocidad en x, $v_x(t)$,
$$v_x(t) = reg(x(t), 2)$$
4. Velocidad en y, $v_y(t)$,
$$v_y(t) = reg(y(t), 2)$$
5. Velocidad absoluta, $v(t)$,
$$v(t) = \sqrt[2]{{v_x}^2(t) + {v_y}^2(t)}$$
6. Aceleración en x, $a_x(t)$
$$a_x(t) = reg(v_x(t), 2)$$
7. Aceleración en y,
8. Aceleración absoluta, $a(t)$,
$$a(t) = \sqrt[2]{{a_x}^2(t) + {a_y}^2(t)}$$
9. Aceleración tangencial, $a_t(t)$,
$$a_t(t) = reg(\|v(t)\|, 2)$$
10. Coseno del ángulo α
$$\cos(\propto) = \frac{v_x(t)}{\|v(t)\|}$$
11. Seno del ángulo α
$$\sin(\propto) = \frac{v_y(t)}{\|v(t)\|}$$
12. Ángulo α
$$\propto = \operatorname{asin}\left(\frac{v_y(t)}{\|v(t)\|}\right)$$
13. Ángulo ϕ
$$\phi(t) = reg(\alpha(t), 2)$$
14. Coseno ángulo ϕ
15. Seno ángulo ϕ
16. Velocidad ángulo ϕ
$$v_\phi = reg(\phi(t), 2)$$
17. Presión $p(t)$
18. Velocidad de *p*, $v_p(t)$
$$v_p(t) = reg(p(t), 2)$$
19. Ángulo azimuth, $az(t)$
20. Velocidad del ángulo azimuth, $v_{az}(t)$

$$v_{az}(t) = reg(az(t), 2)$$

21. Ángulo de inclinación, $in(t)$
22. Velocidad del ángulo de inclinación, $v_{in}(t)$

$$v_{in}(t) = reg(in(t), 2)$$

23. La anchura para una ventana de tamaño 5 centrada en el punto actual.
24. La anchura para una ventana de tamaño 7 centrada en el punto actual.
25. La relación entre la velocidad mínima sobre la máxima para una ventana de 5 puntos, centrada en el punto actual.

Todas las señales anteriores son normalizadas con su desviación estándar:

$$\hat{s} = \frac{s}{\bar{\sigma}_s}$$

4.2. GMM (GAUSSIAN MIXTURE MODELS)[5]

El algoritmo GMM está basado en el modelado de una distribución continua de probabilidad mediante la mezcla de distribuciones de tipo gaussiana.

Una vez que se han extraído las características de la señal tomada del sensor, se hace una comparación con un patrón almacenado en la base de datos de la información. Este proceso da como resultado un score "s" (similitud entre patrón y muestra) que será procesado por un decisor por umbral ("T" o threshold), si "s" es mayor que "T" el sistema decide tomar como cierta la hipótesis ("Él o ella es realmente quien dice ser que es").

La información que se maneja es la misma que la del algoritmo DTW: la posición x, la posición y, la presión, el azimuth y la inclinación.

En el algoritmo GMM se hace primero el procesado que se explicó en el apartado referido al algoritmo DTW para evitar grandes diferencias en tiempo entre las señales a comparar.

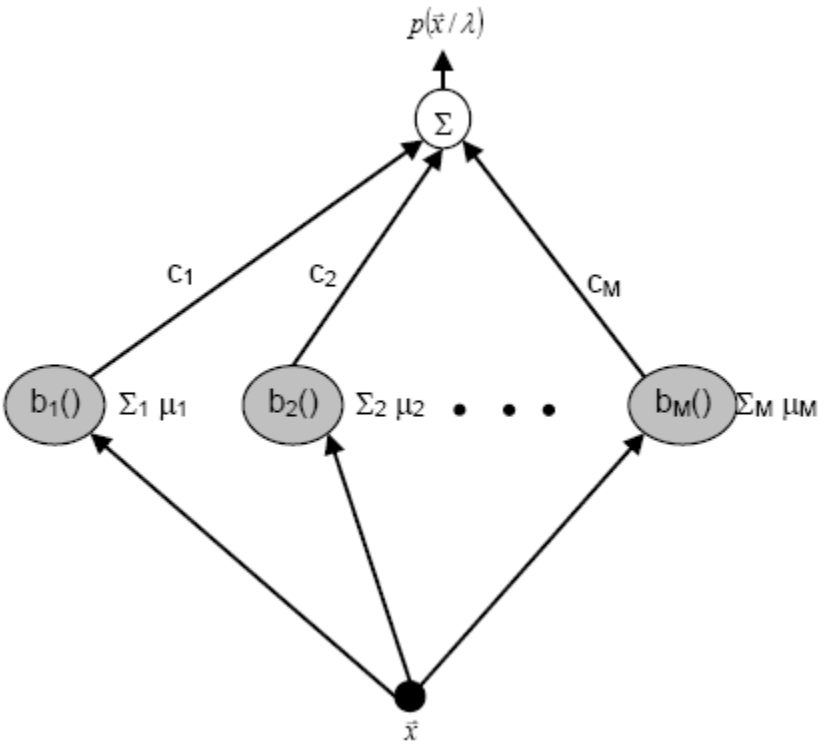


Figura 4.3.: Representación del GMM

[5] O. Miguel-Hurtado, et al., "On-Line Signature Verification by Dynamic Time Warping and Gaussian Mixture Models," in Security Technology, 2007 41st Annual IEEE International Carnahan Conference on, 2007, pp. 23-29.

Según la figura 4.3., a partir de un vector $\vec{x}$, que contiene las características de la firma manuscrita, se obtiene una distribución de probabilidad $p(\vec{x}/\lambda)$, como la combinación lineal entre funciones de densidad de probabilidad de tipo gaussiano.

A continuación se muestran las fórmulas que aplican a este algoritmo:

$$p(\vec{x}/\lambda) = \sum_{i=1}^{M} c_i \cdot b_i(\vec{x})$$

Donde c son los coeficientes aplicados a cada función de densidad de probabilidad que tienen que satisfacer la siguiente condición:

$$\sum_{1}^{M} c_i = 1$$

$b_i()$, son las funciones de densidad de probabilidad que tienen la siguiente forma:

$$b_i(\vec{x}) = \exp\left(-\frac{1}{2} \cdot (\vec{x} - \vec{\mu}_i)' \cdot \Sigma_i^{-1} \cdot (\vec{x} - \vec{\mu}_i)\right) / \left((2\pi)^{L/2} \cdot \|\Sigma_i\|^{1/2}\right)$$

$\vec{\mu}_i$ y Σ_i, son la media y la matriz de covarianzas de cada una de las funciones de densidad de probabilidad.

Cada usuario va a tener su propio modelo en función de los distintos valores de coeficientes c obtenidos. A cada modelo se le denota mediante λ

$$\lambda_s = \{\vec{\mu}_i, \Sigma_i, c_i\} \qquad s = 1 \dots \dots S$$

Donde *S* es el número de usuarios del sistema.

El resultado final de este algoritmo consiste en hacer una comparación de verosimilitudes en función de un umbral de manera que si la relación entre la distribución de probabilidad obtenida de una muestra y la distribución de probabilidad de su patrón es mayor que el umbral “T” se decide como cierta la hipótesis “él o ella es realmente quien dice ser que es”.

5

5. BASES DE DATOS DE FIRMA MANUSCRITA[6]

Los investigadores de Sistemas de Identificación Biométrica han ido recogiendo muestras de distintos rasgos biométricos para poder tener un amplio banco de pruebas que les permita trabajar en el desarrollo de la tecnología biométrica.
Este capítulo se habla de varias de las bases de datos que hay disponibles para el desarrollo de algoritmos de firma manuscrita y que sirven para evaluar los algoritmos que se han descrito en el capítulo 4 de este documento.

[6]J. Ortega-Garcia, et al., "MCYT baseline corpus: a bimodal biometric database," Vision, Image and Signal Processing, IEE Proceedings -, vol. 150, pp. 395-401, 2003.
D. Y. Yeung, et al., "SVC2004: First international signature verification competition," Biometric Authentication, Proceedings, vol. 3072, pp. 16-22, 2004.
B. Dumas, et al., "MyIdea - Multimodal Biometrics Database, Description of Acquisition Protocols," in Third COST 275 Workshop (COST 275), Hatfield (UK), 2005, pp. 59-62.

5.1. INTRODUCCIÓN

La firma es el medio de identificación biométrica con más reconocimiento social y legal. Para un análisis de un sistema biométrico será bueno poseer una base de datos que permita evaluar la eficiencia del mismo. Esta base de datos tendrá que tener distintos usuarios que hagan su firma original y luego se pedirá hacer firmas falsas. De esta manera se sacarán resultados del análisis de estas bases de datos antes de poder sacar al mercado el sistema biométrico. Luego cuando el sistema esté en la calle cada empresa creará sus bases de datos con sus clientes para poder identificarlos.
Para facilitar la labor de los investigadores que puedan evaluar sus procedimientos, tienen a su disposición varias bases de datos. Al tener el banco de pruebas en común se podrá comparar mejor la robustez y eficiencia de los sistemas biométricos basados en firma manuscrita.
En este proyecto se ha hecho uso de tres de estas bases de datos: **SVC2004**, **MCyT** y **MyIDea**. En los siguientes apartados se da una visión acerca de cada una de ellas.

5.2. SVC2004

Para facilitar la evaluación objetiva de los sistemas biométricos basados en firma manuscrita se convoca la Primera Competición Internacional de Verificación de Firma (Signature Verification Competition, SVC) en el año 2003. El objetivo será poner un punto de referencia en las investigaciones que se realizan en este campo de la biometría, estableciendo bases de datos y reglas comunes a seguir por todos los investigadores.

SVC2004 consiste en dos conjuntos de bases de datos de firma manuscrita que se diferencian en la información contenida para cada firma. En el primer conjunto se almacena la información referida a las coordenadas y en el segundo conjunto se añade la orientación del lápiz y la presión. El primer conjunto lo componen 40 usuarios a los cuales se les hacía corresponder 20 firmas originales y 20 firmas falsas. El segundo conjunto lo componen 60 usuarios.

Las firmas se almacenan en ficheros de tipo txt. Se guardará una secuencia de puntos. La primera línea corresponde al número de valores tomadas de la firma. Cada línea posterior corresponde a los siguientes valores (los tres últimos no están en el primer conjunto de la base de datos): coordenadas x, y, tiempo de escritura, estado del lápiz (alzado o no), azimuth, altitud y presión.

Esta base de datos no permite obtener buenos resultados ya que las firmas originales no son realmente originales (por motivos de privacidad a los usuarios les hacían firmar sin su firma), además la firma se hacía con un lápiz sin tinta sobre una tableta gráfica (esto no permitía al usuario ver la firma) y los imitadores disponían de un software de visualización de la dinámica de la firma.

Figura 5.1.: Tableta gráfica

5.3. MCyT

Esta base de datos nace como apoyo para el proyecto MCyT2000: *Aplicación de la Identificación de Personas Mediante Multimodalidad Biométrica en Entornos de Seguridad y Acceso Natural a Servicios de Información*. Se trata de una iniciativa de cuatro universidades españolas cada una de las cuales aportaba algunos usuarios que prestan sus firmas. Estas universidades son:

- Universidad Politécnica de Madrid (UPM). Aporta 145 usuarios a la base de datos.
- Universidad de Valladolid (UVA). Aporta 75 usuarios a la base de datos.
- Universidad del País Vasco/Euskal Herriko Unibertsitatea, Escuela Superior de Ingenieros (EHU). Aporta 75 usuarios a la base de datos.
- Escuela Universitaria Politécnica de Mataró (EUPMT). Aporta 35 usuarios a la base de datos.

Se trata de una base de datos bimodal que contiene la huella dactilar y la firma manuscrita de distintos usuarios. En total se cuenta con 330 usuarios. A cada uno de ellos se le pide 25 firmas originales y también que haga 25 firmas que sean imitación de la firma original otros usuarios.

La forma de adquisición de datos es sobre una tableta gráfica WACOM a la cual se le sobrepone una hoja de papel. La firma es con tinta. Al usuario se le hace firmar en un recuadro.

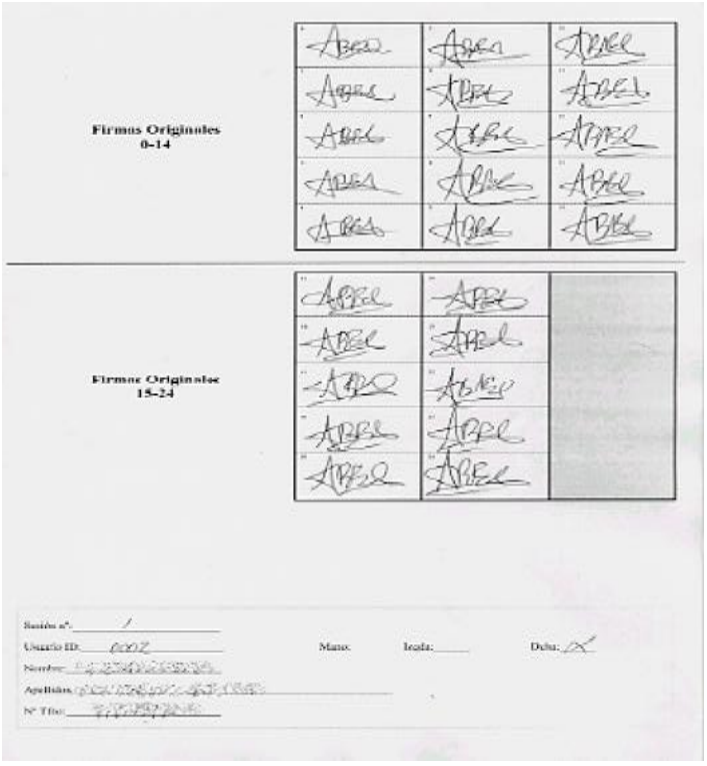


Figura 5.2.: Adquisición de firmas base de datos MCyT

La resolución de la tableta gráfica es de 100 líneas por milímetro y la precisión es de +/- 0.25mm. La frecuencia de las señales tomadas están a 100 Hz. La información que se obtiene de la tableta es la coordenada x, la coordenada y, la presión ejercida sobre el lápiz, el ángulo de azimuth γ_t del lápiz respecto a la tableta y el ángulo de altitud ϕ_t del lápiz respecto a la tableta.

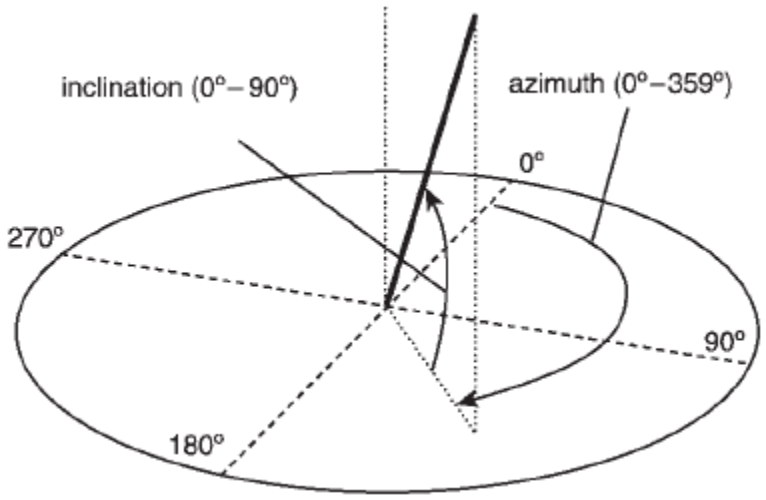


Figura 5.3.: Ángulo de azimuth e inclinación

La toma de las firmas se hace en cinco tandas. En la primera el usuario *n* hace 5 firmas originales suyas y 5 imitando las del usuario *n-1*. Posteriormente volverá a hacer 5 firmas suyas y 5 imitando las del usuario *n-2*. Cuando llegue a imitar las firmas del usuario *n-5* habrá 25 firmas originales del usuario *n* y 25 firmas falsas de distintos usuarios.

5.4.MyIDea

Esta es base de datos multimodal que fue recopilada por la Universidad de Friburgo en Suiza, la Escuela de Ingeniería de Friburgo y el Grupo de Escuelas de Telecomunicaciones (GET) en París.

Esta base de datos surgió en el marco del proyecto nacional suizo IM2 (Interactive Multimodal Information Management) y el proyecto europeo BioSecure.

La base de datos la compone información sobre el rostro, la voz, la huella dactilar, la firma, la escritura, la huella de la palma de la mano y la geometría de la mano. Se tomaron datos de 70 sujetos.

Para tomar la información de la firma se utilizó una tableta WACOM como la usada en las bases de datos anteriores y un bolígrafo de tinta.
La dinámica seguida para grabar las firmas es la siguiente: cada sujeto hace seis firmas genuinas en tres sesiones distintas (18 firmas genuinas por sujeto). Antes de grabar su firma se le pide que entrene un poco para acostumbrarse al lápiz. Para las firmas falsas se pedirá tres imitaciones de sujetos distintos. Habrá dos tipos de imitaciones, estática y dinámica. Para la estática se muestra la forma de la firma original y se pide hacer seis copias (en total el sujeto hará 18 copias, seis por cada imitación). Para la imitación dinámica se muestra también la realización de la firma a través de un programa que repite la firma (*SignReplay*) y también se pedirá hacer seis copias.

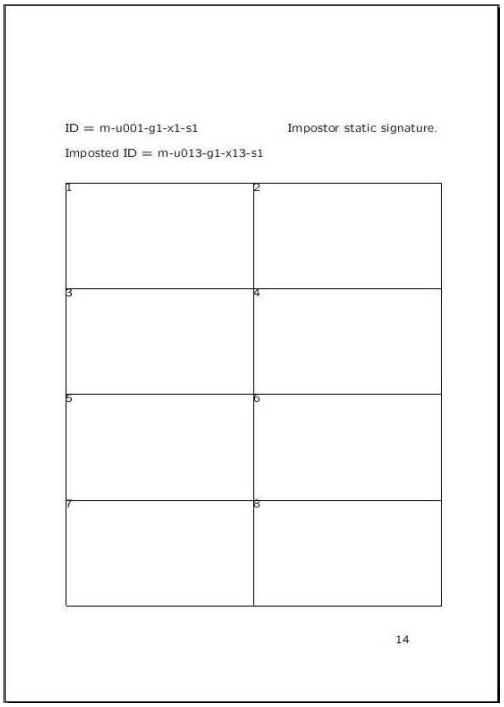


Figura 5.3.: Ejemplo de plantilla para firmas falsas

Al final los datos recogidos en la base de datos correspondientes a la firma manuscrita son los siguientes:

- Firma estática: 630 firmas originales, 3780 imitaciones.

- Firma dinámica: 840 firmas originales, 1260 imitaciones.

(información obtenida de la página web del Departamento de Informática de la Universidad de Friburgo: http://diuf.unifr.ch/diva/biometrics/MyIdea/)

5.5. RESUMEN

En este capítulo se han descrito tres bases de datos que se utilizan en el reconocimiento de firma manuscrita. Existen algunas más como BIOMET, BSEC 2009, BIOSECURE y BiosecurID (estos tres último ofrecen además que no sólo han sido capturadas las firmas mediante tableta gráfica sino también con PDA).

Para las tres bases de datos referidas en el capítulo se ha utilizado una tableta gráfica de la marca WACOM por lo que la información obtenida es prácticamente la misma. La gran diferencia de estas bases de datos se observa en el número de muestras adquiridas para su construcción. A continuación se muestran un par de tablas resumiendo la información que se ha comentado en los apartados anteriores de este capítulo:

Base de datos	**Origen**	**Nº auténticas**	**Nº imitaciones**	**Total/usuario**	**Nº usuarios**	**Total firmas**
MCYT	Occidental	25	25	50	330	16500
SVC2004	Occidental y chinas	20	20	40	40	1600
MyIDea	Occidental	18	36	54	70	3780
	Media/**Total**	*21*	*27*	*48*	**440**	**21880**

Tabla 5.1.: Resumen de nº usuarios por cada base de datos con sus firmas

Base de datos	**Coordenada x**	**Coordenada y**	**Presión p**	**Azimuth γ**	**Elevación ϕ**	**Movimiento lápiz en el aire**
MCYT	X	X	X	X	X	-
SVC2004	X	X	X	X	X	X
MyIDea	X	X	X	X	X	-

Tabla 5.2.: Resumen de información adquirida por cada base de datos

6

6. NORMAS ISO/IEC

La **Organización Internacional para la Normalización** o **ISO** (del griego, ἴσος (*isos*), 'igual', y cuyo nombre en inglés es *International Organization for Standardization*), nacida tras la Segunda Guerra Mundial(23 de febrero de 1947), es el organismo encargado de promover el desarrollo de normas internacionales de fabricación, comercio y comunicación para todas las ramas industriales a excepción de la eléctrica y la electrónica. Su función principal es la de buscar la estandarización de normas de productos y seguridad para las empresas u organizaciones a nivel internacional.

La **Comisión Electrotécnica Internacional** (CEI o *IEC*, por sus siglas del idioma inglés *International Electrotechnical Commission*) es una organización de normalización en los campos eléctrico, electrónico y tecnologías relacionadas. Numerosas normas se desarrollan conjuntamente con la ISO (normas ISO/IEC).

6.1. INTRODUCCIÓN ISO/IEC SC 37

Hemos visto que el uso de la biometría como ciencia para la identificación de personas parte de hace muchos siglos. Quizá a partir de finales del siglo XIX tuvo una mayor importancia en temas forenses y judiciales (trabajos de Bertillon).

Todos estos trabajos no tenían ningún tipo de normalización o estandarización por lo que era imposible la interoperabilidad entres sistemas provenientes de distintos fabricantes.

A finales de los años 90 se empieza a ver la necesidad de crear estándares comunes, así como formatos de datos conocidos y se empieza a trabajar un poco en privado sobre la estandarización biométrica. Fue a partir del atentado del 11 Septiembre de 2001 cuando se pasó a una necesidad mundial. El 20 de agosto de 2002 se crea el Subcomité dedicado a la Identificación Biométrica, dentro del Comité Conjunto ISO/IEC sobre Tecnologías de la Información (JTC1).

El campo de actividad de este Subcomité se centra en: Tecnologías biométricas genéricas, correspondientes a seres humanos, para aportar interoperabilidad e intercambio de datos entre aplicaciones y sistemas. Las normas sobre identificación biométrica humana genérica, incluyen:
Comité Conjunto ISO/IEC sobre Tecnologías de la Información (JTC1).

- Entorno de ficheros comunes: interfaces de programación de aplicaciones biométricas; formato de intercambio de datos biométricos; perfiles relacionados con la biometría; aplicación de criterios de evaluación para las tecnologías biométricas; metodologías para la verificación y emisión de informes sobre el rendimiento de los sistemas, y los aspectos sociales y jurisdiccionales
- De todo este trabajo se excluye:
    - El trabajo realizado dentro del ISO/IEC JTC1/SC17, sobre la aplicación de las tecnologías biométricas a las tarjetas y a la identificación personal
    - El trabajo realizado dentro del ISO/IEC JTC1/SC27, sobre las técnicas de protección de los datos biométricos, la verificación de la seguridad biométrica, sus evaluaciones y sus metodologías de evaluación.

Se crean seis grupos de trabajo cada uno encargado de trabajar en un aspecto determinado del la Identificación Biométrica:

- WG1, *Harmonised Biometric Vocabulary*: encargado de hacer un catálogo con los términos estandarizados de un sistema biométrico.
- WG2, *Biometric Technical Interfaces*: trabajan sobre los interfaces de comunicación entre aplicaciones y sistemas.
- WG3, *Biometric Data Interchange Formats*: trabajan sobre los formatos que deben cumplir los datos para cada modalidad biométrica.
- WG4, *Biometric Functional Architecture & Related Profiles*: definen las distintas arquitecturas biométricas y sus perfiles de aplicación.

- WG5, *Biometric Testing and Reporting*: encargado de definir los mecanismos de evaluación así como los formatos de los informes de resultados.
- WG6, *Cross-jurisdictional and Society Aspects*: estudia los efectos jurídicos de la instalación de sistemas biométricos y la influencia social de la aplicación de los mismos.

6.2. PROYECTO 19794

El proyecto 19794 tiene que ver con la definición de los formatos de los datos de las distintas técnicas biométricas. Se encuadra por tanto dentro del grupo de trabajo WG3 que se mencionó en el apartado anterior.

El proyecto 19794 dispone de las siguientes partes:

- ISO/IEC 19794-1: 2006 Information Technology. Formato de datos biométricos para el intercambio entre aplicaciones. Marco general que resume los demás trabajos.
- ISO/IEC 19794-10: 2007 Information Technology. Formato de los datos para sistemas biométricos basados en la geometría de la mano.
- ISO/IEC 19794-2: 2005 Information Technology. Formato de los datos para sistemas biométricos basados en minucias de los dedos.
- ISO/IEC 19794-3: 2005 Information Technology. Formato de los datos para sistemas biométricos basados en patrones de dedo.
- ISO/IEC 19794-4: 2005 Information Technology. Formato de los datos para sistemas biométricos basados en imágenes de dedo.
- ISO/IEC 19794-5: 2005 + A2: 2009 Information Technology. Formato de los datos para sistemas biométricos basados en imágenes del rostro.
- ISO/IEC 19794-6: 2005 Information Technology. Formato de los datos para sistemas biométricos basados en imágenes de iris.
- ISO/IEC 19794-7: 2007 Information Technology. Formato de los datos para sistemas biométricos basados en la firma manuscrita.
- ISO/IEC 19794-8: 2008 Information Technology. Formato de los datos para sistemas biométricos basados en el esqueleto del dedo.
- ISO/IEC 19794-9: 2007 Information Technology. Formato de los datos para sistemas biométricos basados en imágenes vasculares
- ISO/IEC 19794-11, Information Technology. Formato de los datos para sistemas biométricos basados en firmas dinámicas
- ISO/IEC 19794-13, Information Technology. Formato de los datos para sistemas biométricos basados en la voz
- ISO/IEC 19794-14, Information Technology. Formato de los datos para sistemas biométricos basados en el ADN

Algunos de estos estándares todavía están sin publicar, en fase de escritura.

6.3. FORMATO DE DATOS 19794-7 FULL FORMAT[7]

La serie 19794-7 trata sobre el formato de los datos recogidos de las firmas. Para ello se utilizan tabletas gráficas o lápices especiales que permitan recoger información digital.

Para la aplicación de los formatos que se definen será necesario también tener conocimiento de los siguientes documentos:

- ISO/IEC 19785-1, Information Technology. Parte referida a la especificación de los elementos de los datos.
- ISO/IEC 19785-2, Information Technology. Parte referida a los procesos de operaciones para la Biometric Registration Authority.
- ISO/IEC 19785-3, Information Technology. Parte referida a la especificación de formatos de patrones.
- ISO/IEC 19794-1, Information Technology. Formato de datos biométricos. Parte referida a la estructura.

Para grabar los datos se distinguen dos tipos de formatos: Full Format o Compact. Como se puede deducir en un dato Full Format la información guardada es más amplia. A continuación se muestran dos esquemas explicativos de los dos formatos.

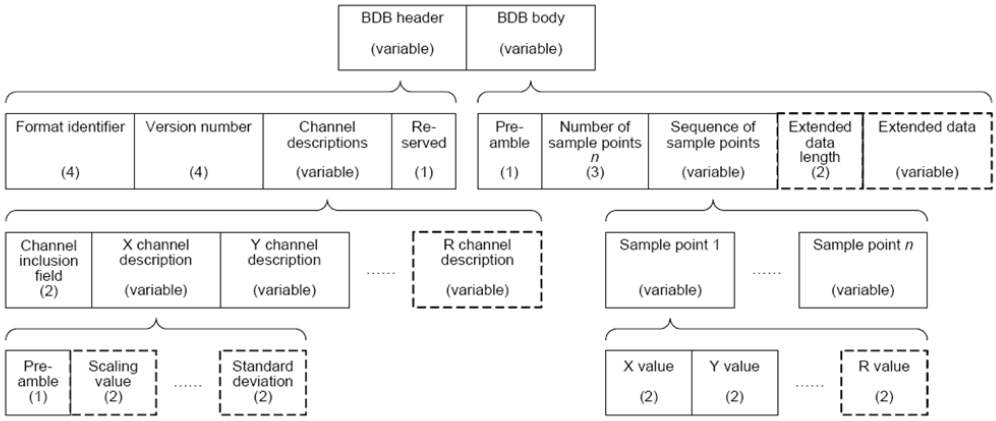


Figura 6.1.: Formato de un bloque de datos Full Format

[7] O. Miguel-Hurtado, et al., "Analysis on compact data formats for the performance of handwritten signature biometrics," in Security Technology, 2009. 43rd Annual 2009 International Carnahan Conference on, 2009, pp. 339-346.
Information technology — Biometric data interchange formats —Part 7:Signature/sign time series data ISO/IEC FDIS 19794-7,2006

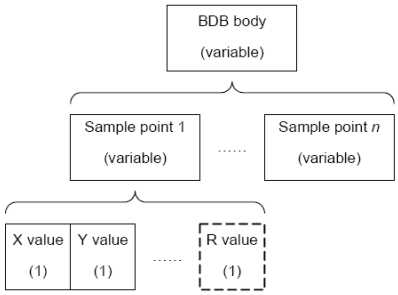


Figura 6.2.: Formato de un bloque de datos Compact

Para este proyecto se ha usado el formato Full Format que a continuación se analiza un poco más en detalle:

6.3.1. BDB Header

Este bloque guarda información global de la firma.

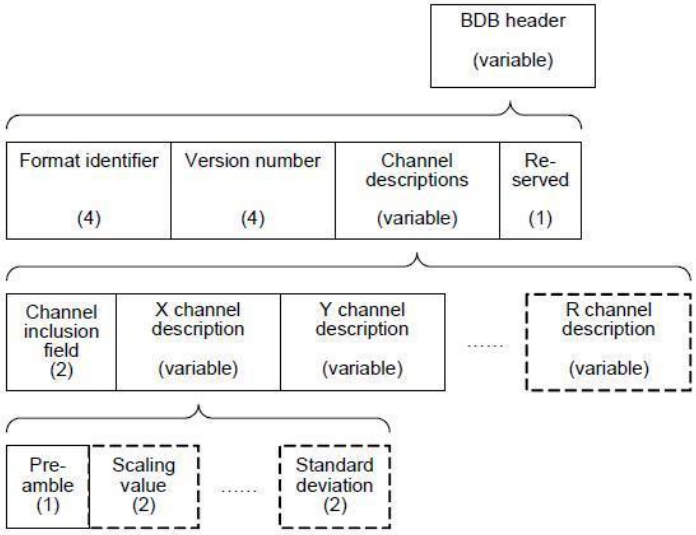


Figura 6.3. : Formato cabecera Full Format

Esta parte consta de 4 partes que se describen en la siguiente tabla.

BDB Header	**Número de octetos**
Format identifier	4
Version number	4
Channel descriptions	Variable
Reserved	1

- Format identifier. Este campo consiste en los tres caracteres ASCII “SDI” seguidos de un octeto nulo. De esta manera se sabe que el bloque de datos corresponde a una firma.

- Version number: son 4 bytes que identifican la versión del formato que se está utilizando.
- Channel descriptions: Los datos que incluyen este campo indican que información ha sido capturada y almacenada en el cuerpo de la muestra biométrica y cuáles son sus principales características. Está formado por los siguientes campos:

    - Channel inclusion field. Este campo lo forma dos bytes de manera que cada bit corresponde a distintos canales que pueden incluirse en la captación de los datos de una firma. La siguiente tabla posiciona los distintos canales que se pueden incluir. Si el bit correspondiente vale "1" esto quiere decir que el canal está incluido, si es "0" esto es que no se incluye. Los canales X e Y son obligatorios, mientras que también tienes que existir al menos uno de los 3 siguientes indicares de la frecuencia de muestreo: T o DT o información sobre muestreo constante. Los demás canales son opcionales.

Nombre del canal	**Octeto**	**Posición de bit**
X	1	8(MSB)
Y	1	7
Z	1	6
VX	1	5
VY	1	4
AX	1	3
AY	1	2
T	1	1(LSB)
DT	2	8(MSB)
F	2	7
S	2	6
TX	2	5
TY	2	4
Az	2	3
El	2	2
R	2	1(LSB)

    - Channel description preamble. Esto es un byte y cada bit indica si se incluye una información (si vale "1") o si no se incluye (si vale "0"). En la siguiente tabla se muestra la información que se puede incluir. A este campo le seguirán los valores que se hayan indicado como existentes en este Channel description preamble para cada canal.

Atributo del canal	**Posición de bit**
Valor de escala	8(MSB)

Valor mínimo posible del canal	7
Valor máximo posible del canal	6
Valor medio de los valores	5
Desviación estándar de los valores	4
Valor constante	3
Componente lineal	2
RFU(reservado para uso futuro)	1(LSB)

6.3.2. BDB Body

En estos bytes se guarda la información que se ha descrito en el campo Header que se ha descrito en el apartado anterior.

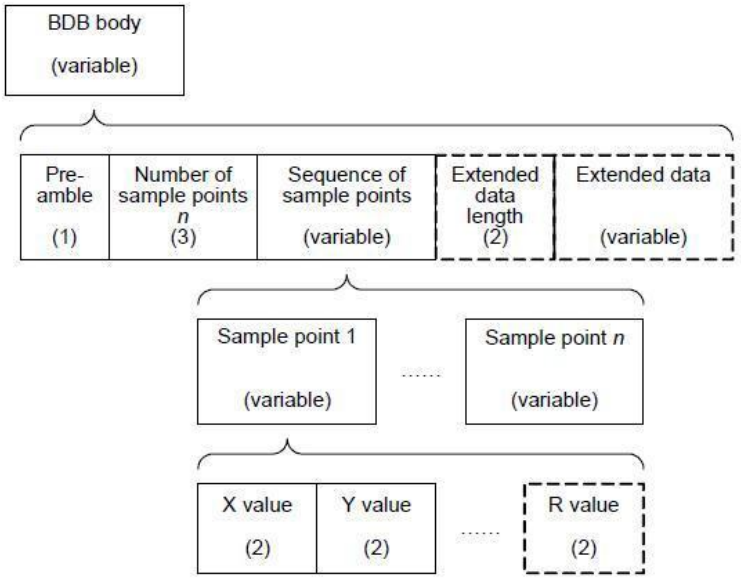


Figura 6.4.: Formato body Full Format

Este campo comienza con un octeto que indica la presencia o no de datos extendidos.
A continuación se indica el número de muestras almacenadas codificado en tres octetos. A partir de este punto comenzará la secuencia de muestras que variará en función del número de canales incluidos y que se codificaron en la cabecera.

En función de los valores de los bytes del campo “Channel inclusion field” los valores de las muestras varían al incluir o no los distintos canales.
Los valores de los canales Z, T, DT, F, Az, El y R están comprendidos entre 0 y65535 codificados en dos octetos como enteros sin signo.
Los valores de los canales X, Y, VX, VY, AX, AY, TX y TY están comprendidos entre -32768 y 32767 codificados en dos octetos como enteros sin signo después de sumar a cada valor 32768.
El canal S tendrá el valor 0 o 1 codificado en un octeto como entero sin signo.

6.4. PROYECTO 19795

El proyecto 19795 tiene que ver con la definición de los mecanismos de evaluación de los sistemas biométricos así como los formatos de los informes de resultados.

Este proyecto consta de 7 partes:

- ISO/IEC 19795-1: 2006 Information Technology. Principios y estructura.
- ISO/IEC 19795-2: 2007 Information Technology. Metodología de ensayo para la tecnología y escenario de evaluación.
- ISO/IEC 19795-3: 2007 Information Technology. Modalidad específica de ensayo.
- ISO/IEC 19795-4: 2008 Information Technology. Interoperabilidad de funciones de ensayo.
- ISO/IEC 19795-5, Information Technology. Esquema de puntuación para un sistema de control de acceso.
- ISO/IEC 19795-6, Information Technology. Metodología de ensayo para la evaluación operacional.
- ISO/IEC 19795-7, Information Technology. Ensayo para la comparación de algoritmos biométricos.

Las partes en las que no aparece indicado el año, están todavía en desarrollo.

A la hora de diseñar un Sistema de Identificación Biométrica se busca que sea seguro ya que va a servir como un medio de autenticación y además va a utilizar datos muy sensibles de cada individuo. Por tanto es importante ver el rendimiento funcional del Sistema Biométrico para saber si es más o menos fiable. El trabajo realizado por el proyecto 19795 expone unas bases para evaluar los Sistemas Biométricos siguiendo unos métodos y normas que permitirán compararlos y observar el rendimiento o la eficiencia de los mismos.

7

7. EVALUACIÓN[8]

Todo trabajo realizado debería ser evaluado antes de sacarlo al mercado. Se trata de ver los posibles errores que se pueden cometer y actuar en consecuencia. El objetivo final sería obtener un producto fiable y competente. Para la evaluación se pueden probar distintas características del producto viendo los resultados obtenidos. Estas pruebas pueden ser desarrolladas por el investigador o puede utilizar bancos de pruebas de trabajos anteriores.
En este capítulo se va a presentar la evaluación en el campo de la firma manuscrita.

[8] A.J. Mansfield and J.L.Wayman, "Best Practices in Testing and Reporting Performance of Biometric Devices", Aug.2002
Information Technology—Biometric performance testing and reporting—Part 1: Principles and Framework, ISO Standard ISO/IEC FDIS 19795-1, 2005
Information Technology—Biometric performance testing and reporting—Part 2: Testing Methodologies for Technology and Scenario Evaluation, ISO Standard ISO/IEC FDIS 19795-2, 2006

7.1. INTRODUCCIÓN A LA EVALUACIÓN

En el capítulo anterior se ha hablado del proyecto 19795 realizado por el Subcomité Internacional de ISO ISO/IEC/JTC1/SC37 con el cual se pretende establecer un marco común para la evaluación de los Sistemas Biométricos. Algunos de los trabajos realizados dentro de este proyecto ya han sido publicados y sirven como referencia a los investigadores. Otros trabajos todavía están en proceso de realización.

Hay más textos que también hablan acerca de la evaluación de los Sistemas de Identificación Biométrica. Destacar el *"Best practices in testing and reporting performance of biometric devices"* producido por el Grupo de Trabajo de Biometría (Biometric Working Group). En el sumario de este documento se exponen los objetivos:

- Proveer un esquema para el desarrollo y la total descripción de protocolos de pruebas.
- Ayudar a evitar desviaciones sistemáticas debido a colecciones de datos incorrectos o procedimientos erróneos de evaluación.
- Ayudar a las pruebas a obtener las mejores estimaciones posibles del rendimiento gastando el mínimo esfuerzo al realizar sus evaluaciones.
- Mejorar en el entendimiento de los límites de aplicabilidad de los resultados de las pruebas y los métodos de las pruebas.

Estos objetivos resumen lo que se pretende conseguir con el trabajo realizado en la normalización y estandarización de la evaluación de los Sistemas de Identificación Biométrica. Hay luego más factores también muy importantes que deberían ser analizados en los Sistemas de Identificación Biométrica pero que ni en este documento ni en los del proyecto 19795 se tratan. Estos son:

- Realizabilidad, disponibilidad y mantenimiento.
- Vulnerabilidad
- Seguridad
- Aceptación social
- Factores humanos
- Coste/beneficio
- Conformidad a la regulación privada.

Estos aspectos se deberán analizar en cada Sistema de forma individual antes de sacarlo al mercado.

En la siguiente figura se muestra un esquema general de la evaluación de un Sistema de Identificación Biométrica.

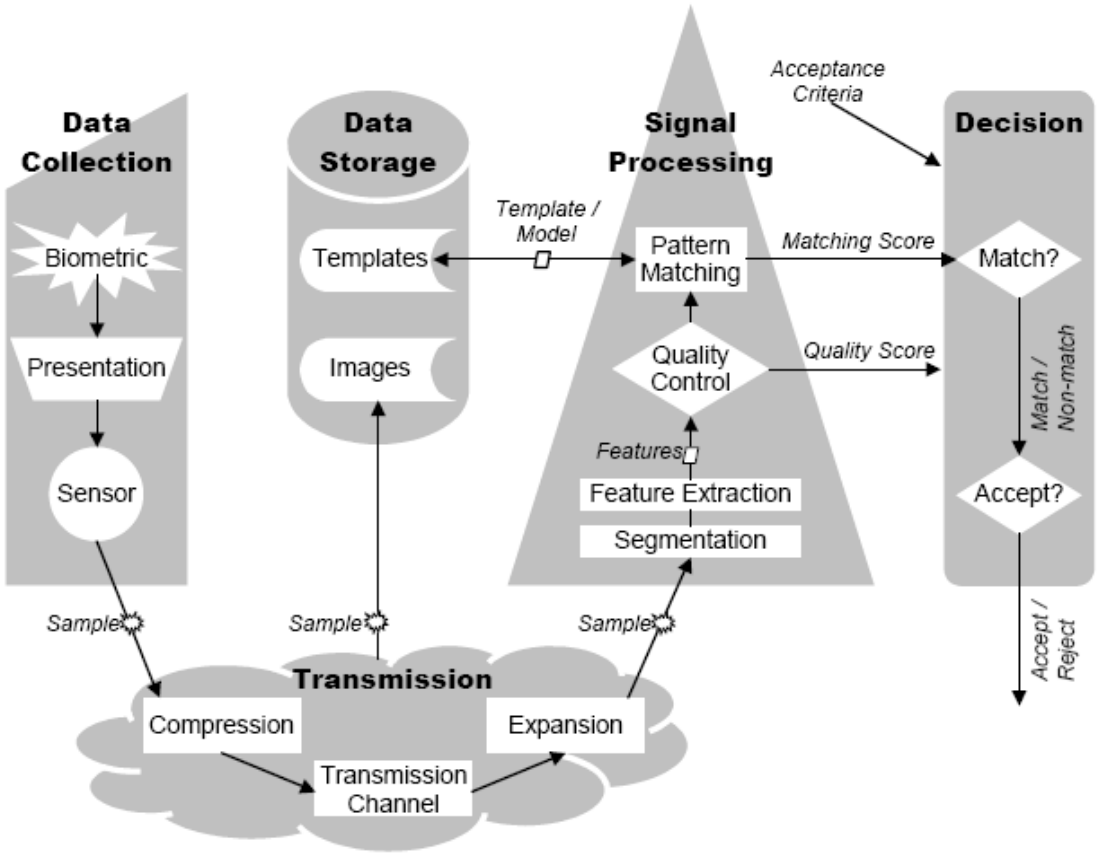


Figura 7.1.: Esquema general de un Sistema de Identificación Biométrica

El primer bloque “DATA COLLECTION” corresponde a los componentes físicos del Sistema. En el caso de la firma manuscrita el sensor será el instrumento que permita digitalizar una firma como puede ser una tableta gráfica o un lápiz digitalizador. El sensor genera una muestra (sample, si se utiliza nomenclatura anglosajona).
Esta muestra pasa al bloque “TRANSMISSION” donde se comprime y tras pasar por el canal de transmisión que corresponda pasará a guardarse en una base de datos (“DATA STORAGE”) o si la muestra es solicitada para una comparación se expande para pasar al bloque “SIGNAL PROCESSING”.
El bloque “SIGNAL PROCESSING” realiza el cálculo de las similitudes de las muestras. Antes se hacen unas plantillas o modelos (templates) que se almacenan también en “DATA STORAGE”. Las muestras que sean evaluadas se comparan con el modelo que le corresponda para obtener su similitud.
Finalmente este valor de similitud pasa a la etapa “DECISION” en la cual se acepta o rechaza la muestra analizada.

7.2. TIPOS DE ERRORES

En el apartado anterior vimos el esquema de evaluación del Sistema de Identificación Biométrico. Se distinguen cinco etapas. A la hora de evaluar el sistema existirán distintos errores en función de en qué etapa se esté considerando.

ERRORES EN LA TOMA DE DATOS (DATA COLLECTION)

- **FTER (Failure to Enroll Rate)**: es la tasa entre las altas insatisfactorias y el número total de intentos de darse de alta.
- **TFI (Tasa de Fallos en Inscripción)**: proporción de muestras no inscritas por no cumplir algún criterio requerido.
- **TFO (Tasa de Fallos en Operación)**: se estima mediante la proporción de operaciones (clientes e impostores) que no han podido completarse.

ERRORES EN LA CLASIFICACIÓN DE DATOS (SIGNAL PROCESSING)

- **TFP o FMR (Tasa de Falsos Positivos)**: proporción de muestras de usuarios diferentes clasificadas como el mismo.
- **TFN o FMNR (Tasa de Falsos Negativos)**: proporción de muestras del mismo cliente clasificadas como diferentes usuarios.

ERRORES EN LA ALMACENACIÓN DE DATOS (DATA STORAGE)

- **Coeficiente de penetración**: es la proporción de la base de datos total que se examina, en promedio por cada muestra de entrada.
- **Error de clasificación**: refleja los errores inducidos por las inconsistencias en el proceso de partición de los patrones y de clasificación de las muestras.

ERRORES EN LA DECISIÓN FINAL (DECISION)

- **TFA o FAR (Tasa de Falsas Aceptaciones)**: es la proporción esperada de operaciones con identidad falsamente reclamada que son incorrectamente confirmadas.
- **TFR o FRR (Tasa de Falsos Rechazos)**: es la proporción esperada de operaciones con identidad o no identidad correctamente reclamada que son incorrectamente rechazadas.

En el proyecto ya contamos con las bases de datos construidas por lo que nos centramos en la etapa “DECISION” para evaluar el rendimiento de los algoritmos de identificación de firma manuscrita. Las distintas bases de datos se utilizan para generar los modelos o patrones variando el número de firmas utilizadas y que luego se comparan con el resto de las firmas para comprobar la variación de los errores en la etapa de decisión.

7.3. REPRESENTACIÓN GRÁFICA

En el apartado anterior se ha hecho referencia a algunos de los errores que se pueden observar en la evaluación de un Sistema de Identificación Biométrico. Algunos de ellos serán más importantes que otros y también admitirán una representación gráfica que facilitará el análisis de los resultados obtenidos. En este proyecto sólo nos fijamos en los errores FAR y FRR que van a ser suficientes para evaluar la eficiencia de los distintos algoritmos de identificación de firma manuscrita que se usan.

A continuación se muestran algunas gráficas que se pueden incluir en los informes de evaluación.

FAR-FRR

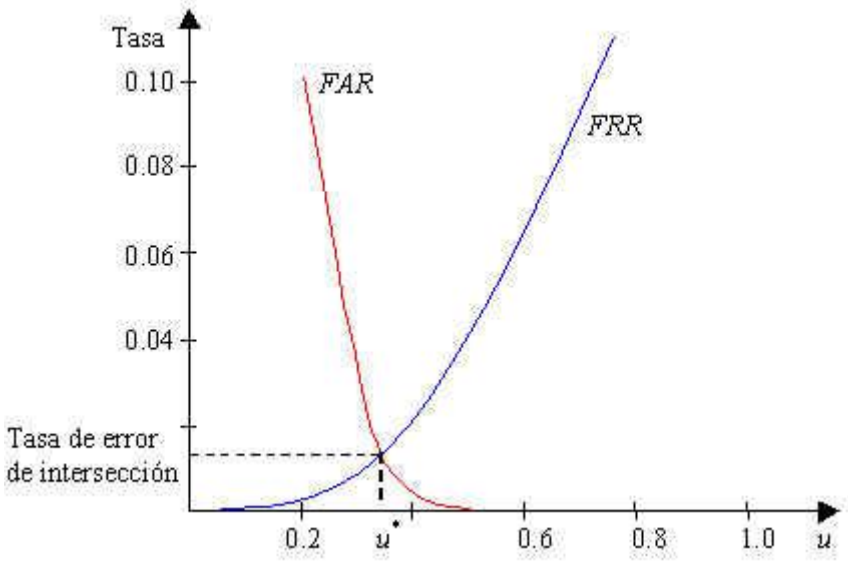


Figura 7.2.: Gráfica FAR y FRR

En esta gráfica se observa el error de las Falsas Aceptaciones y el de los Falsos Rechazos. Como sabemos estos errores corresponden a la fase de decisión, de ahí que en el eje X aparezca el umbral respecto al cual se decide. Para la construcción de esta gráfica se va variando el umbral y se cuentan las Falsas Aceptaciones y los Falsos Rechazos en proporción al número total de experimentos realizados de identificación.
El punto de intersección entre las dos curvas se conoce como Tasa de error de intersección (EER: equal error rate, en sus siglas en inglés). Interesa que este valor sea pequeño para que el sistema sea más preciso.
Tener FAR con valores altos implica que el sistema identifica erróneamente muchas veces (los imitadores son aceptados).
Tener FRR con valores altos implica que el sistema no identifica correctamente muchas veces (los clientes son denegados).
Habrá que llegar a un punto de compromiso para que las dos tasas sean aceptables. Normalmente se toma como valor umbral el EER, aunque esto puede variar

dependiendo del uso que se le vaya a dar al sistema, y del nivel de seguridad que se quiera obtener.

DET(Detection Error Tradeoff)

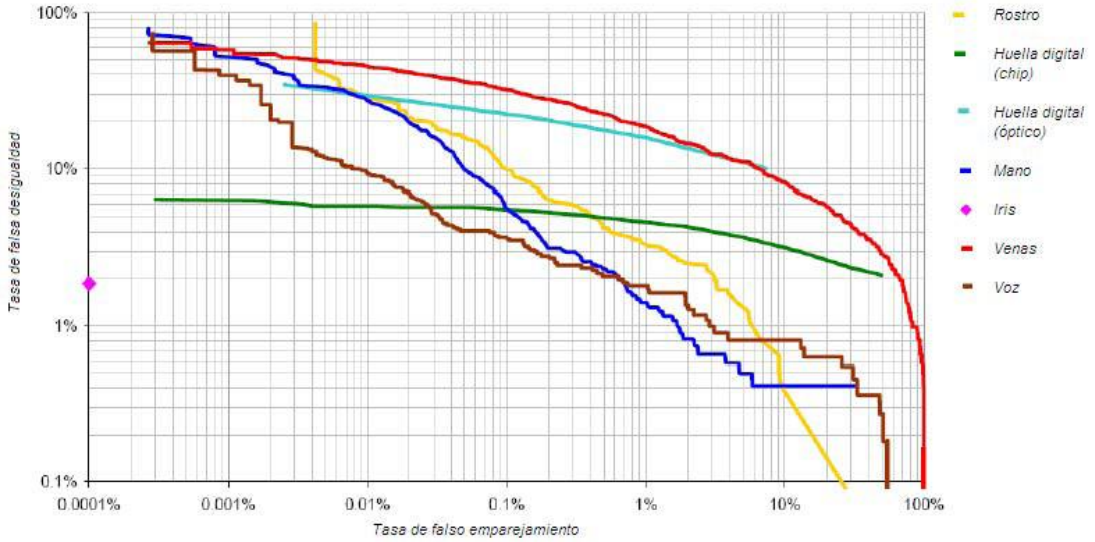


Figura 7.3.: Curvas DET para varios rasgos biométricos

Las curvas DET trazan las tasas de error de coincidencias (tasa de falsos positivos o FMR vs tasa de falos negativos o FNMR).
Esta curva permite ver el comportamiento del error del sistema con base en las distribuciones de porcentajes de acierto y error observados.

ROC (Receiver Operating Characteristics)

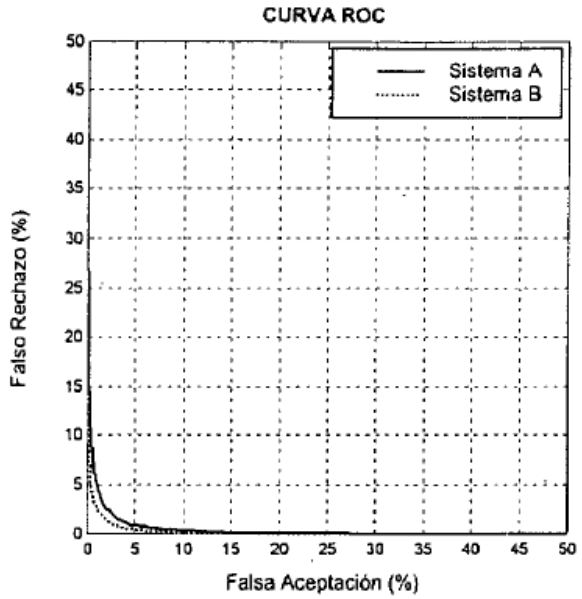


Figura 7.4.: Curva ROC de un sistema basado en huella dactilar

La curva ROC es un método aceptado para resumir el funcionamiento del Sistema de Identificación Biométrica. Permite hacer un diagnóstico de la detección y el emparejamiento de patrones. Es un gráfico que muestra cómo varía la FRR en función de la FAR. Es decir ROC es una función dibujada sobre unos ejes de coordenadas cartesianas cuyo eje de coordenadas es FRR y cuyo eje de abscisas es FAR. Las curvas ROC son independientes del umbral por lo que permiten una comparación del funcionamiento de sistemas diferentes en condiciones similares.

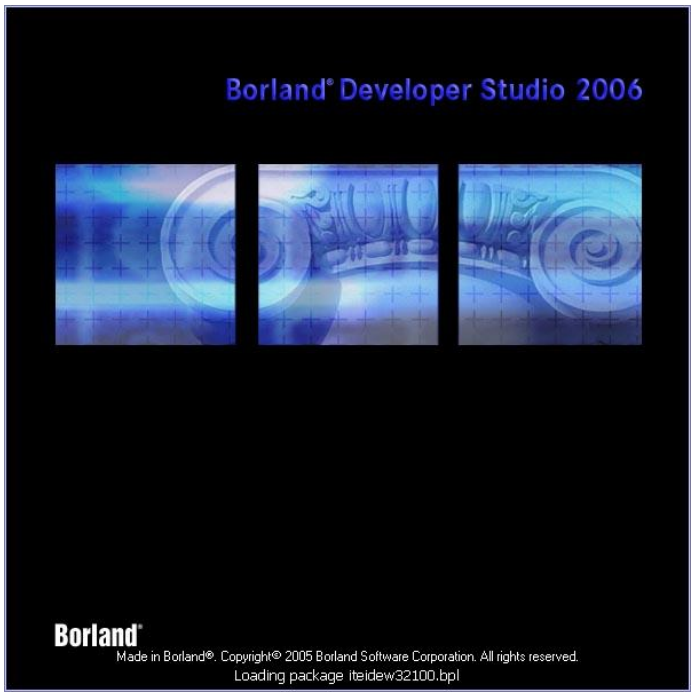
Estas gráficas son las más utilizadas en la evaluación de los Sistemas Biométricos. Se podría combinar los distintos errores para obtener nuevas gráficas pero para ver la eficiencia sirve con ver las curvas FAR-FRR y la ROC. En la realización de este proyecto esas dos gráficas son las que se utilizan porque permiten comparar distintos algoritmos de identificación de firma manuscrita variando distintos parámetros de simulación de las bases de datos utilizadas como puede ser el número de intervalos de umbral utilizados o el número de firmas utilizadas para generar el patrón.

8

8. DISEÑO, ARQUITECTURA Y DESARROLLO

En el capítulo 1 de este documento se hablaba de que para la realización de la interfaz se ha utilizado como herramienta el Borland Developer Studio 2006 que nos permite trabajar con C++ realizando Programación Orientada a Objetos (P.O.O.).

La interfaz consiste en distintas ventanas que se muestran en función de operación que se solicita.



8.1. DISEÑO Y FUNCIONALIDADES DE LA APLICACIÓN

La interfaz gráfica que se ha diseñado tiene por objetivo evaluar bases de datos de firmas manuscritas con los distintos algoritmos que se hayan programado.

El siguiente esquema muestra las distintas funcionalidades de la aplicación programada.

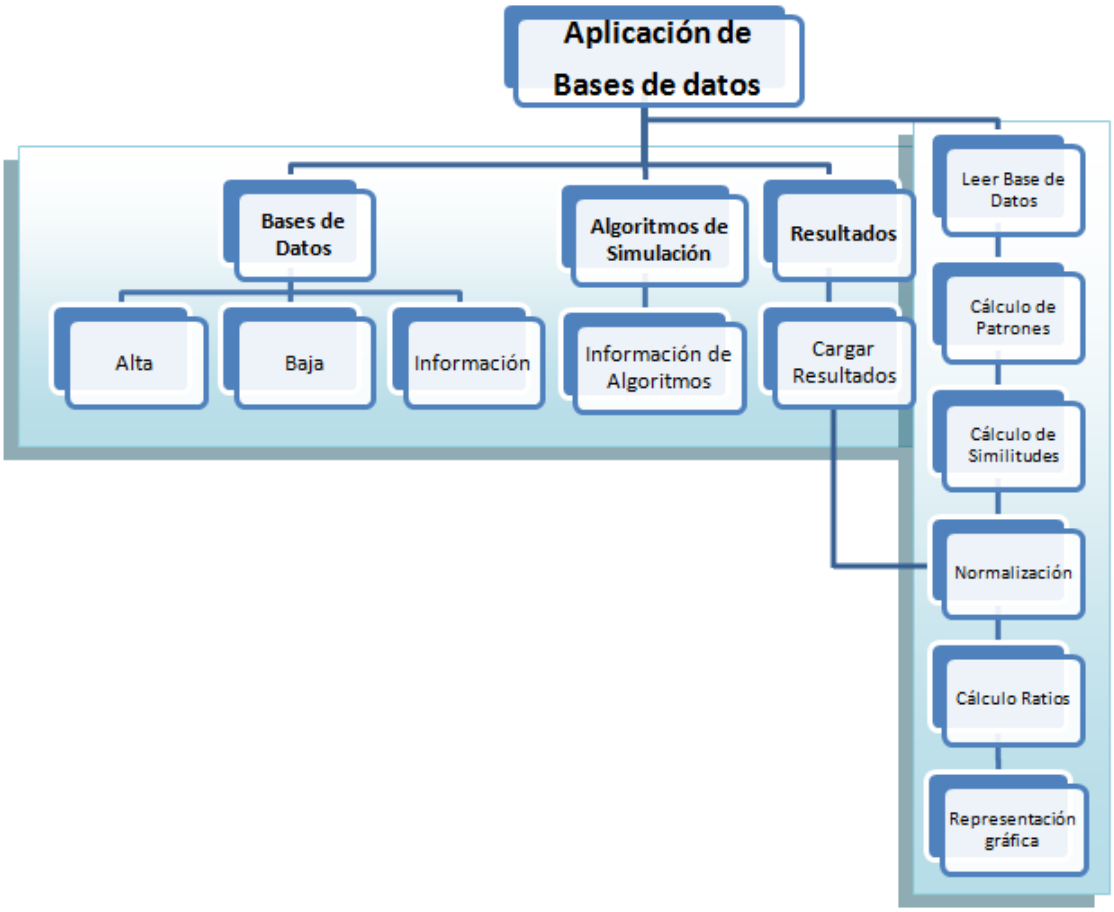


Figura 8.1.: Esquema de la Aplicación

En el esquema anterior, el recuadro grande de la izquierda encuadra a las distintas funciones de los menús que se encuentran en la ventana principal de la aplicación. El recuadro de la derecha corresponde a los procesos que se realizan a la hora de realizar una simulación.

En los siguientes apartados de este capítulo se analiza cada proceso.

8.2. APLICACIÓN DE BASES DE DATOS

Al ejecutar la aplicación se dará de alta a todas las bases de datos que hubiera disponibles de aplicaciones anteriores. Dar de alta significará decir el nombre, la ruta, número de usuarios y número de firmas tanto originales como falsas. Esta información se obtiene de un fichero de texto que por defecto se llama "*bbdd.txt*" y que está en la carpeta "*C:\\BBDD*" (esto se podrá modificar en el código).
Las bases de datos se almacenan en memoria siguiendo una estructura de una lista enlazada, de manera que cada elemento de la lista corresponde a una base de datos que apuntará al siguiente elemento. Cada nodo de la lista tiene guardado la información referente al nombre, ruta, información, número de usuarios y número de firmas originales y falsas, de una base de datos.

La estructura del fichero "*bbdd.txt*" será la siguiente:

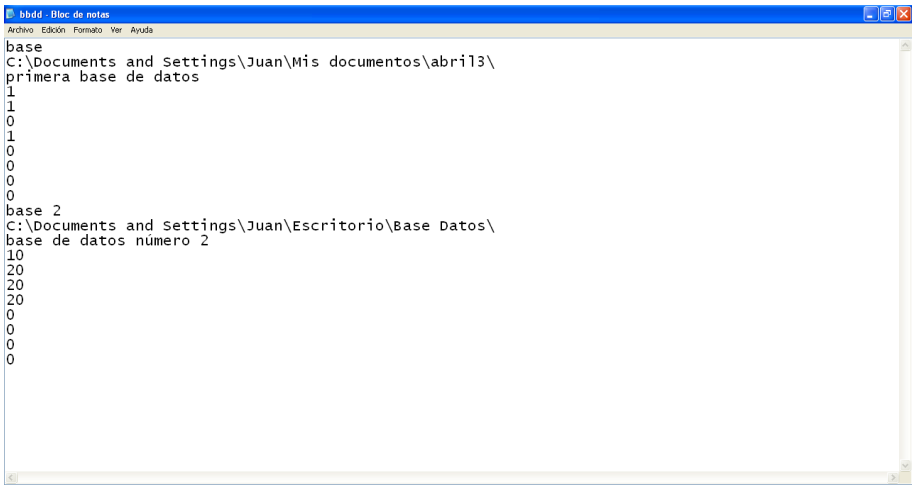


Figura 8.2.: "*bbdd.txt*"

La primera línea del fichero corresponde al nombre de la base de datos, a continuación se pone la ruta donde está almacenada, la siguiente línea muestra información relativa a la base de datos, las siguientes líneas corresponden a los datos numéricos como número de usuarios, número de firmas genuinas y número de firmas falsas (de los distintos tipos desde el 1 hasta el 6). En la figura anterior se muestran dos bases de datos distintas.

Una vez que se ha dado de alta a todas las bases de datos que hubiera en el fichero "*bbdd.txt*", se muestra la ventana principal de la aplicación.

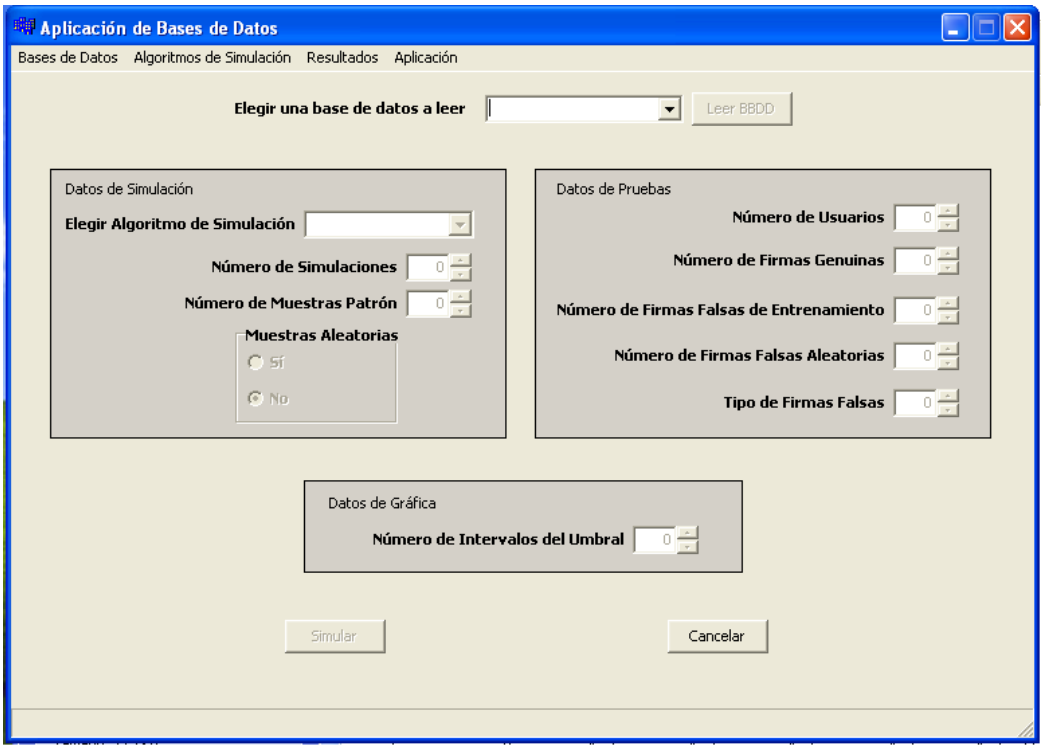


Figura 8.3.: Ventana principal de la aplicación

Por defecto todos los campos aparecen deshabilitados a excepción del campo de "Elegir una base de datos a leer" y los distintos menús. La lectura de una base de datos se hace en el formato ISO 19794-7 Full Standard que ya fue descrito en el capítulo 6 de esta memoria. Cuando se lee una base de datos lo que se obtiene es un puntero a estructuras de tipo BDB_FullFormat que lo que contienen son los distintos campos que forman una firma y que ya se comentó al describir el formato ISO 19794-7.

8.3. BASES DE DATOS

Este menú trata la gestión de todas las bases de datos que se manejan durante el uso de la aplicación. Como se dijo anteriormente puede haber algunas bases de datos que ya se carguen al iniciar el programa y habrá otras que se puedan añadir, también se podrán quitar o mostrar la información relativa a las distintas bases de datos. Son entonces tres operaciones las que se pueden hacer con las bases de datos, cada una de las cuales tiene sus ventanas con los campos necesarios para su realización. A continuación se muestra cada operación con su ventana correspondiente y una pequeña descripción.

- Alta:

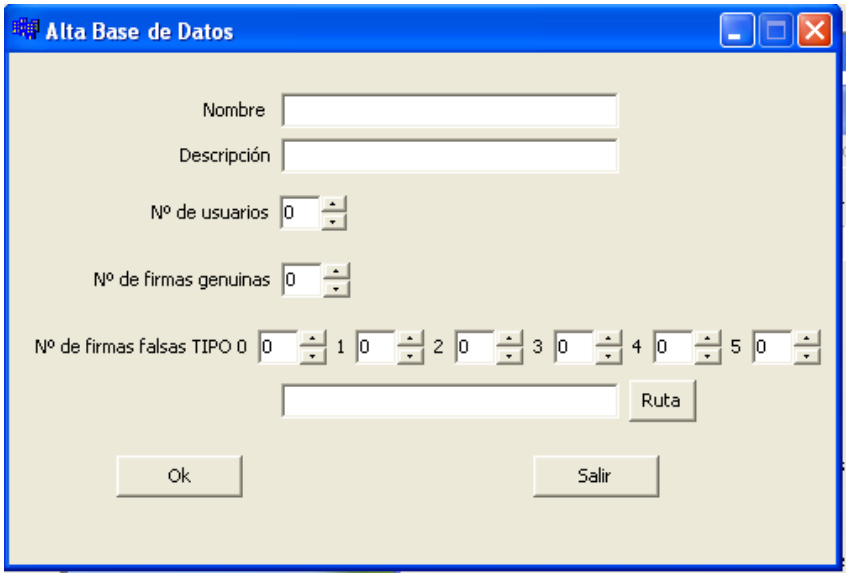


Figura 8.4.: Ventana de menú Alta

En esta ventana se rellenan los distintos campos y al pulsar el botón “Ok” se comprueba que el nombre de la base de datos no coincida con el de otra que ya esté en memoria. También se comprueba que los datos numéricos sean coherentes. Para buscar la ruta hay otra ventana que aparece cuando se pulsa el botón “Ruta”.

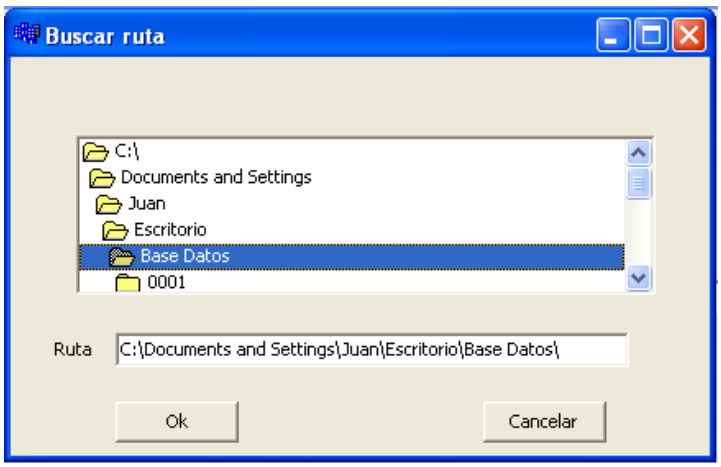


Figura 8.5.: Ventana de ruta para Alta de Base de Datos

Esta ventana muestra la estructura raíz de los directorios. Cuando se selecciona un directorio aparece escrito en “Ruta” y se habilita entonces el botón “Ok” que al ser pulsado cierra la ventana copiando “Ruta” a la ventana de Alta.

- Baja:

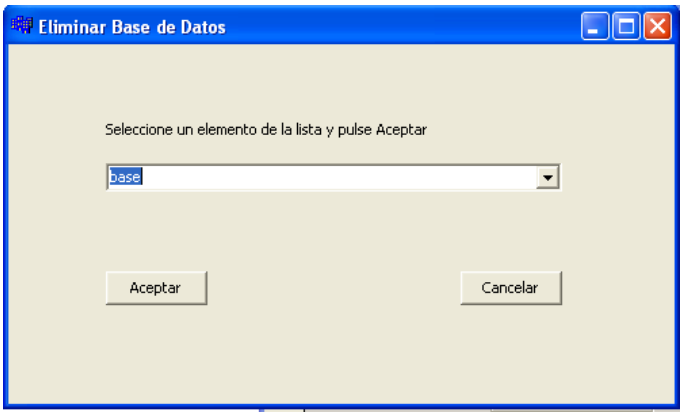


Figura 8.6.: Ventana de menú Baja

Esta ventana muestra una lista con todas las bases de datos que están dadas de alta. Al seleccionar una base de datos y pulsar el botón “Aceptar”, ésta se elimina de la lista enlazada. Si no hay una base de datos seleccionada el botón “Aceptar” está deshabilitado.
Esta función se usa para cuando alguna base de datos ya no quiere ser utilizada en la evaluación de un algoritmo. De esta manera se evita que pueda ser seleccionada una base de datos que ya no existe o que ha cambiado de ubicación. En el caso de que se cambie la ruta de una base de datos lo que hay que hacer es dar de baja primero y a continuación dar de alta con la nueva ruta.

- Información:

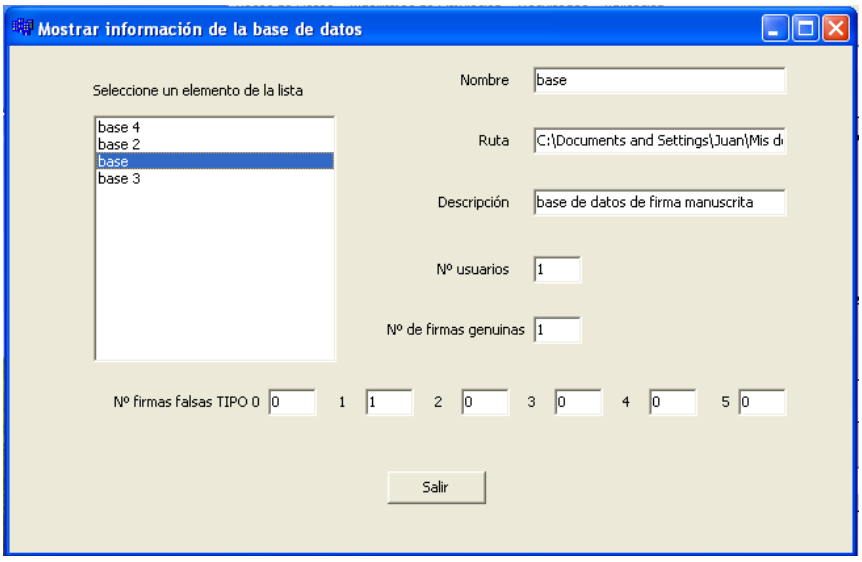


Figura 8.7.: Ventana menú Información

En esta ventana se muestra una lista con todas las bases de datos disponibles en la aplicación y al seleccionar una de ellas se rellenan todos los campos con su información.
Al seleccionar un elemento de la lista de la izquierda se procede a buscar esa posición en la lista enlazada para sacar entonces la información relativa a esa base de datos. Esto nos puede servir principalmente para ver la capacidad de las distintas bases de datos disponibles en la aplicación y de esta manera se puede ver qué campos se pueden modificar en la simulación del algoritmo.

8.4. ALGORITMOS DE SIMULACIÓN

Este menú sirve para mostrar los algoritmos de simulación disponibles en el programa. Se pueden añadir más algoritmos según la necesidad del investigador. En el caso del proyecto que nos ocupa, sólo se necesitan los dos algoritmos que se describieron en el capítulo 4 de esta memoria, que son el GMM y el DTW.

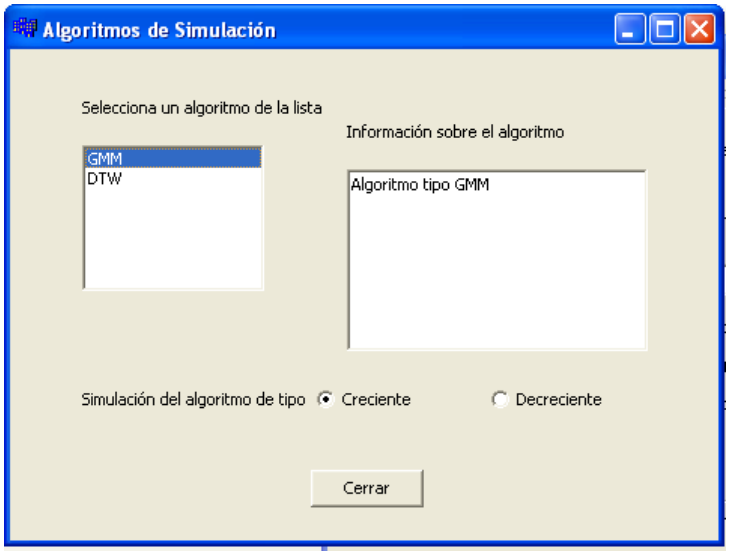


Figura 8.8.: Ventana menú Algoritmos de Simulación

En esta ventana se muestran los algoritmos de simulación disponibles en la aplicación. Al seleccionar un algoritmo se da una información sobre él y se indica el tipo de simulación correspondiente.

8.5. RESULTADOS

En algunos casos se puede haber guardado simulaciones que se hayan realizado anteriormente. De esta manera se podría ahorrar la lectura de la misma base de datos. También facilitaría la realización de informes sobre el algoritmo. Se pueden hacer distintas simulaciones variando los parámetros y cada resultado de la simulación se guarda en ficheros distintos. Posteriormente se pueden sacar esos resultados sin necesidad de realizar de nuevo la simulación. Los ficheros son de tipo binario y para poder cargarlos tienen que tener por nombre "resultados_*" donde * podrá ser cualquier nombre, pero el comienzo del nombre de fichero tiene que ser el mismo.

El fichero va a almacenar datos de la simulación como usuarios, número de firmas genuinas, número de firmas falsas, número de firmas falsas aleatorias y sobre todo va a guardar los datos de las similitudes calculadas a la hora de hacer la simulación.

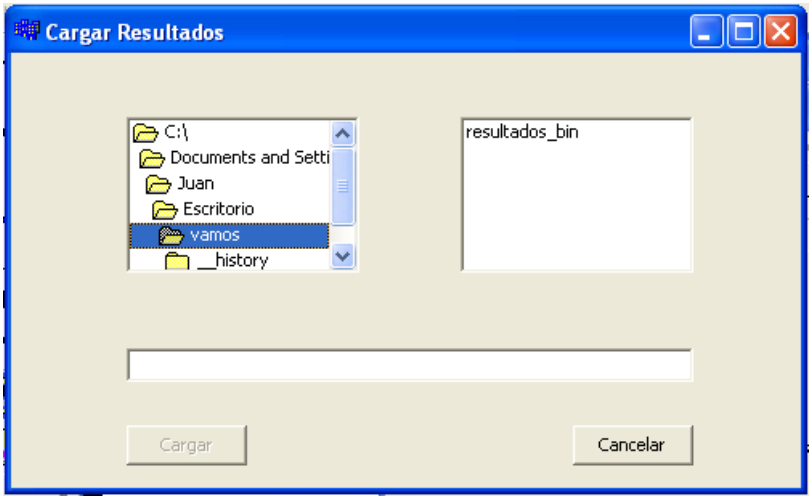


Figura 8.9.: Ventana menú Resultados

En esta ventana se elige un fichero tipo binario que tenga guardados datos referidos a antiguas simulaciones. Una vez que se le diera al botón "Cargar" se mostraría la ventana con las gráficas de la simulación.

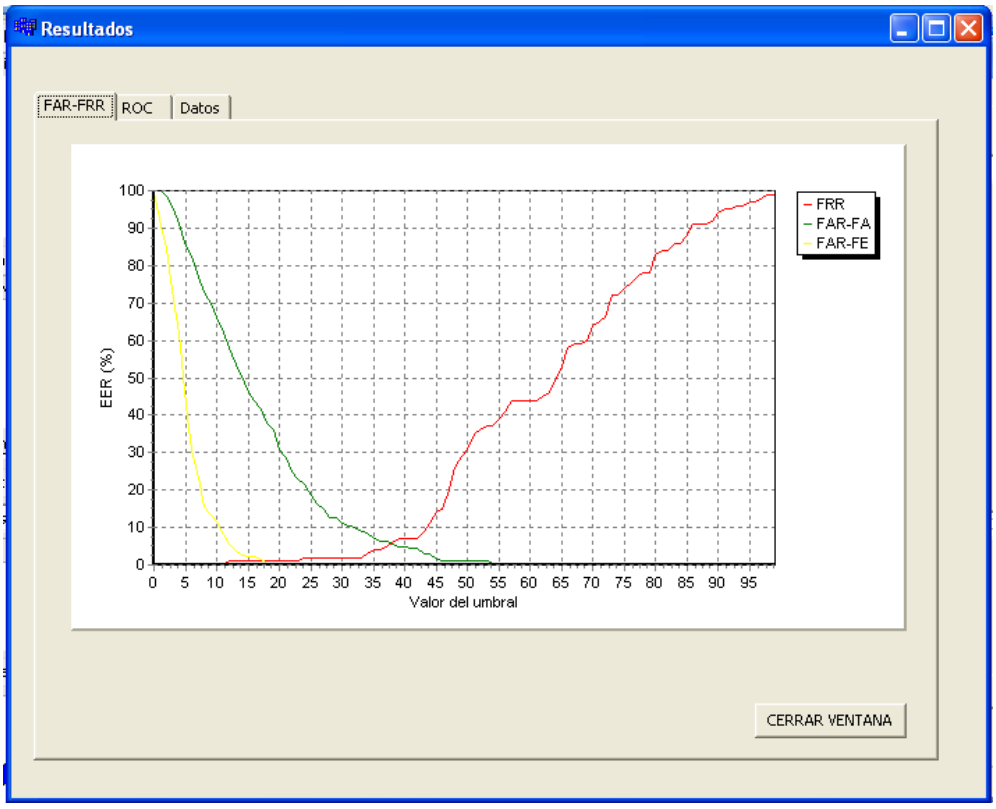


Figura 8.10.: Ventana de resultados gráfica FAR-FRR

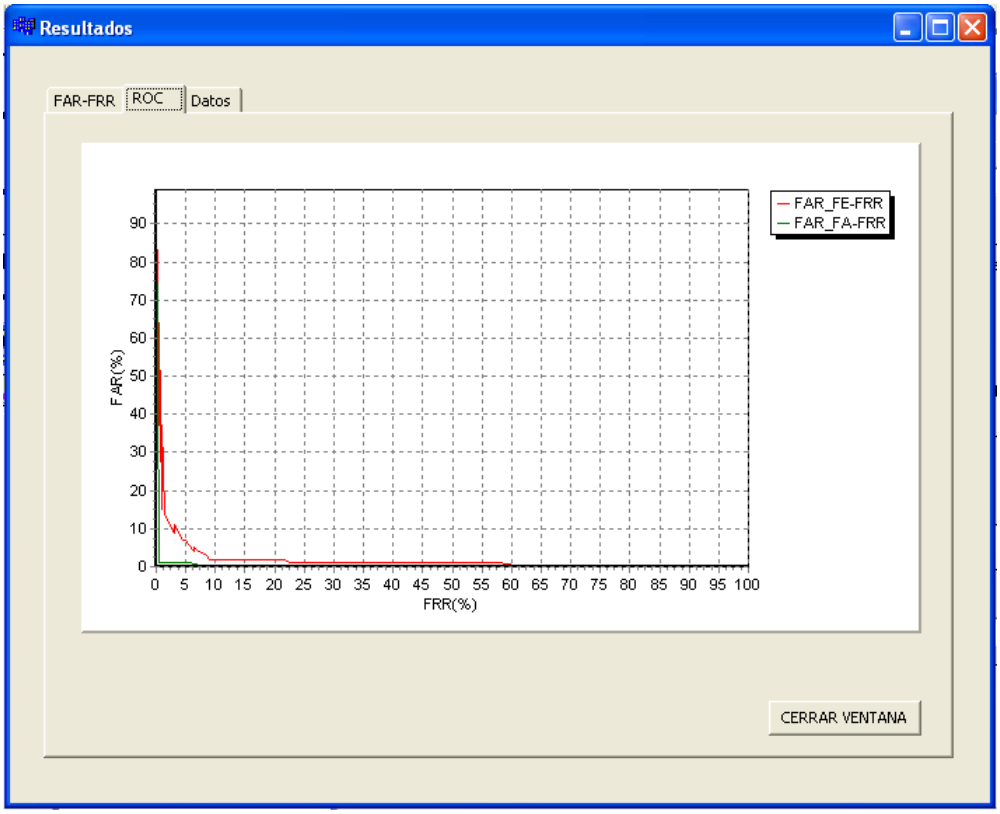


Figura 8.11.: Ventana de resultados gráfica ROC

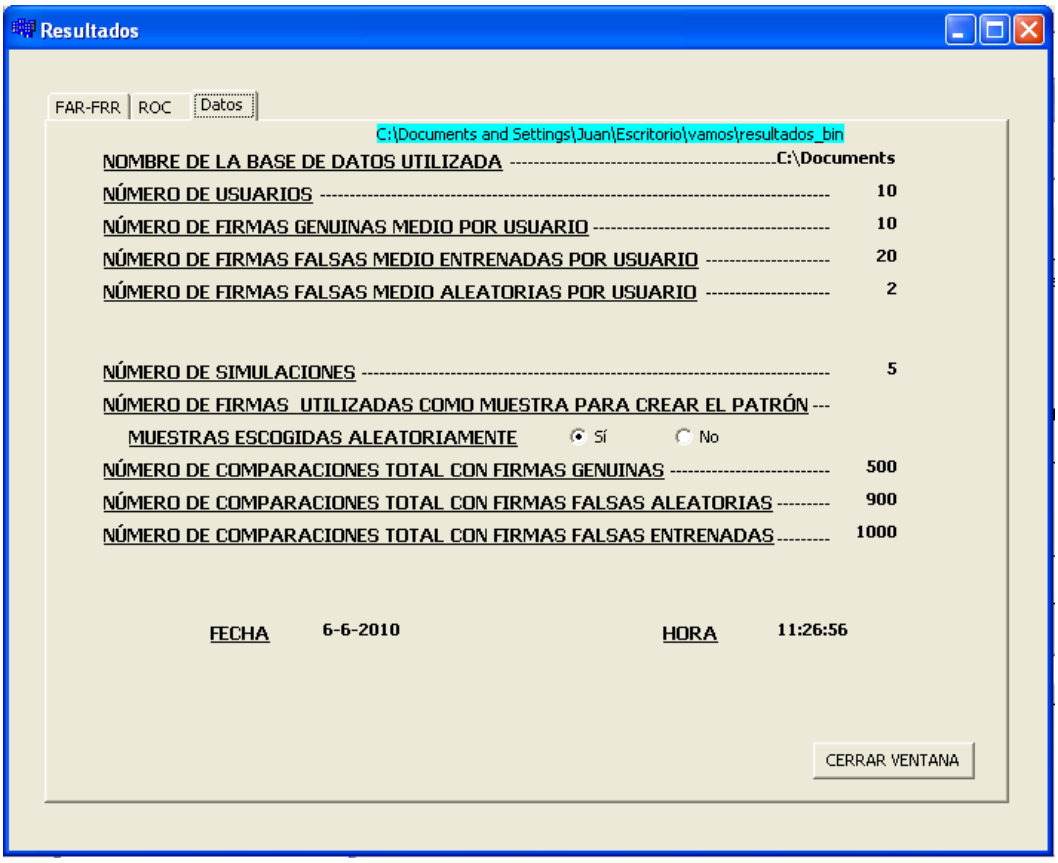


Figura 8.12.: Ventana de resultados datos de simulación

La ventana de resultados muestra dos gráficas distintas, colocadas en dos pestañas. Una es la FAR-FRR y la otra es la ROC. Estas dos gráficas se describieron en el capítulo 7 de la memoria. Se comentaba que estas gráficas nos sirven para evaluar la eficiencia de los algoritmos de identificación de firma manuscrita. En la última pestaña se muestran los datos referidos a la simulación.

8.6. SIMULACION DE UNA BASE DE DATOS

8.6.1. LECTURA BASE DE DATOS

En la figura 8.3. se muestra la ventana principal donde hay un campo para seleccionar una base de datos que esté en la lista enlazada de la aplicación. Una vez que se selecciona una de esas bases de datos y se pulsa el botón “Leer bbdd” se procede a la lectura.

Cada nodo de la base de datos contiene información como su nombre, la ruta que indica dónde está guardada, y el número de firmas originales y falsas de distintos tipos que contiene.

Cuando se lee una base de datos se llama a una función que va generando los nombres de los ficheros que se van a leer dentro de un bucle que recorra todos los usuarios y todas las firmas. En el nombre la primera parte corresponde a la ruta y lo que va variando es el nombre del fichero.
La base de datos tiene que tener tantas carpetas como usuarios, cada una con el nombre “000X” siendo X el número de usuario. Luego dentro de cada carpeta hay ficheros tipo “.ssf” cada uno de los cuales corresponde a una firma. El nombre de ese fichero tiene que tener el mismo formato:

- Firmas genuinas: “U000X_S00Y_G_T0.ssf”, donde X es el número de usuario, Y es el número de firma, G es de firma genuina y el tipo será T0.
- Firmas falsas: “U000X_S00Y_F_TZ.ssf”, donde X es el número de usuario, Y es el número de firma, F es de firma falsa y Z es el tipo de firma falsa.

Para la lectura de cada fichero se almacena en memoria una estructura “Muestra” de datos en C que contiene todos los datos de la firma (ver capítulo 6.3.Formatos de datos 19794-7 Full Format).

Cuando se ha leído la base de datos, en memoria hay un vector de estructuras tipo “Usuario” de datos en C, en el que para cada usuario tiene almacenadas las “Muestras” de cada firma genuina y falsa.

La aplicación está ya en condiciones para pasar al siguiente paso para la evaluación de la base de datos.

8.6.2. CÁLCULO DE PATRONES

Una vez que se ha leído la base de datos hay que sacar algunas características de las firmas de cada usuario para poder comparar las firmas y decidir. Se cogen varias firmas genuinas para generar un puntero que corresponde a un patrón del usuario.

En la ventana principal de la aplicación se indica el número de muestras usadas por cada usuario para calcular los patrones de la firma. Se toman las muestras genuinas de cada usuario para generar un patrón por cada uno de ellos. El número de muestras que se usen para tal propósito influirá a la hora de conseguir un modelo más robusto. Si se cogen pocas muestras el modelo nos dará un mayor porcentaje de errores porque las características obtenidas del patrón no serán consistentes. Para realizar la simulación tampoco será conveniente utilizar todas las muestras genuinas de cada usuario porque entonces no quedarán muestras para observar los resultados obtenidos al evaluar las muestras genuinas con el algoritmo.

Las muestras usadas para calcular los patrones no se utilizan posteriormente en la simulación ya que éstas serán siempre aceptadas.

8.6.3. CÁLCULO DE SIMILITUDES

En este proceso se va a hacer corresponder a cada muestra una similitud. Para ello se recorren todas las muestras genuinas y falsas y en función del patrón que corresponda a cada usuario, se le asigna un valor numérico.

Para el cálculo de las similitudes se utilizan las firmas genuinas no usadas en el cálculo de los patrones. También se usan firmas falsas entrenadas correspondientes a imitaciones de otros usuarios y también se pueden coger aleatoriamente firmas de otros usuarios tanto genuinas como falsas. De esta manera se puede evaluar mejor el funcionamiento del algoritmo viendo su comportamiento ante un posible ataque de fuerza bruta por ejemplo (al introducir una firma aleatoria). El número de firmas falsas entrenadas y aleatorias puede ser indicado en la ventana principal de la aplicación.

En este proyecto se ha incluido la posibilidad de tener seis tipos de firmas falsas distintas:

- Firma Falsa Tipo 0: firma de un usuario falsificador que no ha visto la firma original.
- Firma Falsa Tipo 1: firma de un usuario falsificador que ha visto la firma original ya hecha. Sólo tiene un intento para hacer la falsificación.
- Firma Falsa Tipo 2: firma de un usuario falsificador que ha visto la firma original ya hecha. Puede entrenar para hacer la falsificación.
- Firma Falsa Tipo 3: firma de un usuario falsificador que ha visto cómo se realizaba la firma genuina. Sólo tiene un intento para hacer la falsificación.
- Firma Falsa Tipo 4: firma de un usuario falsificador que ha visto cómo se realizaba la firma genuina. Puede entrenar para hacer la falsificación.
- Firma False Tipo 5: firma de un usuario falsificador que tiene varias imágenes de firmas genuinas en distintos documentos.

Al final se obtiene una matriz de valores tipo float con todas las similitudes de todos los usuarios y todas sus muestras. Estos valores se normalizan para que no se obtengan valores muy dispersos. La normalización consiste en restar a cada valor el valor mínimo total y dividirlo entre la diferencia del valor máximo y el mínimo.

8.6.4. SIMULACIONES

Cuando ya se tiene la matriz de similitudes de una base de datos se procede a comparar estos valores con los distintos valores de umbral. En la ventana principal se puede indicar el número de veces que se va a realizar la simulación y el número de intervalos en los que se divide el umbral (son valores entre 0 y 100).

Se realizan tantas repeticiones como se indiquen en la ventana principal. De esta manera se consiguen resultados más reales de las comparaciones ya que para cada simulación se toman las muestras de manera aleatoria y así se consigue que los resultados no dependan de las muestras utilizadas para generar el patrón.

Se hace un bucle que recorre todos los valores de similitudes obtenidos para compararlos con cada valor de umbral. De esta manera se va completando un vector que lleva la cuenta de los falsos aceptados y los falsos rechazos. Una vez que se han completado los vectores se hace la media dividiendo entre el número total de firmas genuinas o falsas y multiplicando el valor por 100.

Se obtienen tres vectores en los que se almacenan los ratios de falsos aceptados con firmas falsas entrenadas, falsos aceptados con firmas falsas aleatorias y falsos rechazos con firmas genuinas. Estos son los errores FAR (False Aceptance Rate) y FRR (False Rejection Rate).

8.6.5. ANÁLISIS DE RESULTADOS

Una vez que se han realizado las simulaciones se procede a representar gráficamente los resultados obtenidos de las falsas aceptaciones y los falsos rechazos.

Se muestra una ventana con tres pestañas correspondientes a dos gráficas y un cuadro con los datos de la simulación. En la primera de las gráficas se muestra la variación del error FAR (False Acceptance Rate) y el error FRR (False Rejection Rate) en función de los valores del umbral. La segunda de las gráficas corresponde a la ROC (Receiver Operating Characteristic) que representa el error FAR frente al error FRR.

Mediante esta ventana se podrá observar la variación de los errores del algoritmo frente a los cambios en el umbral de decisión. Ayudará a los investigadores a escoger el mejor valor para cada algoritmo llegando a un compromiso entre los dos errores FAR y FRR.

En la ventana también se podrá dar la opción de guardar los datos de la simulación para una posible consulta posterior y de esta manera ahorrar el trabajo de lectura y análisis de la misma base de datos (esta opción no es facilitada cuando se ha cargado una simulación previa).

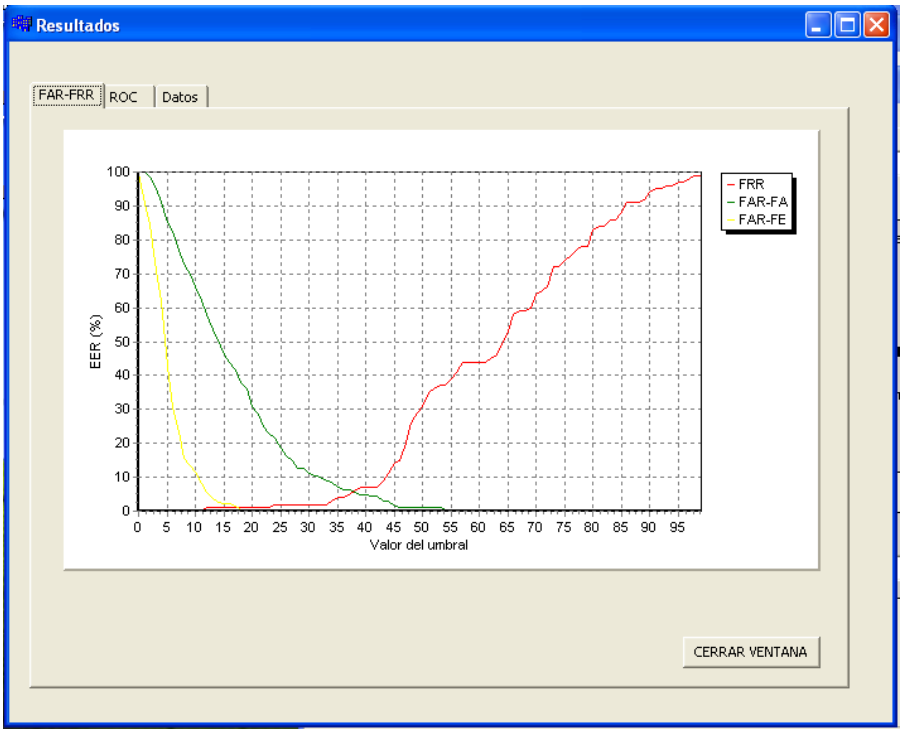


Figura 8.13.: Resultados de gráfica FAR-FRR de una simulación

En esta ventana se muestran los resultados de los ratios que se describieron en el apartado anterior. En la leyenda se ve FRR para los falsos rechazos con firma manuscrita, FAR-FE para los falsos aceptados con firmas falsas entrenadas y FAR-FA para los falsos aceptados con firmas falsas aleatorias.

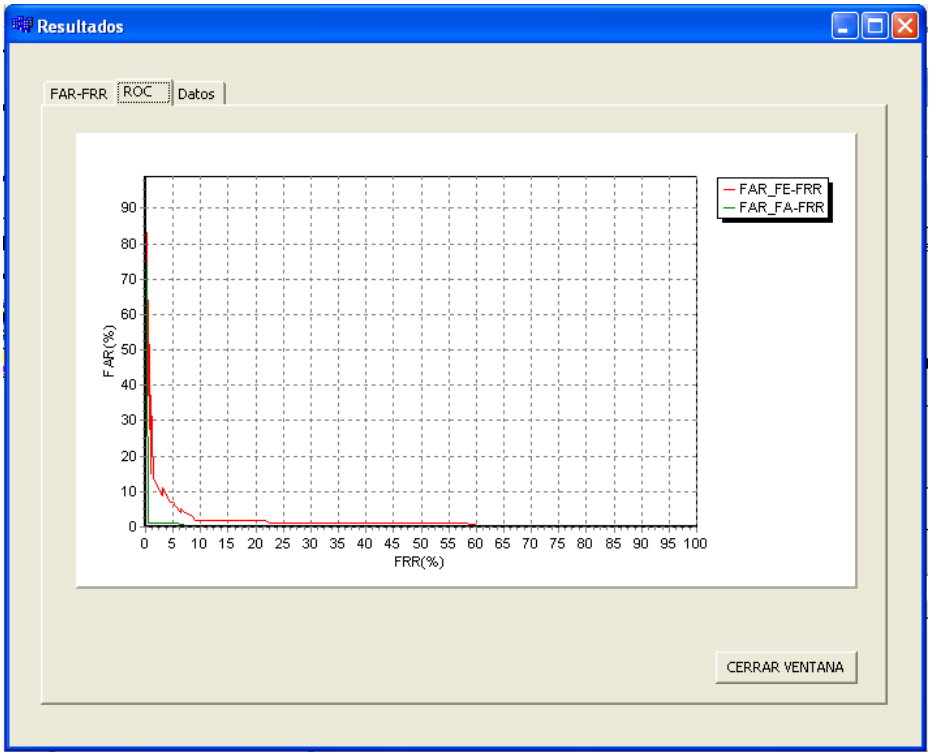


Figura 8.14.: Resultados de gráfica ROC de una simulación

Esta gráfica representa los valores de FAR frente a los FRR tanto de firmas falsas entrenadas como de firmas falsas aleatorias.

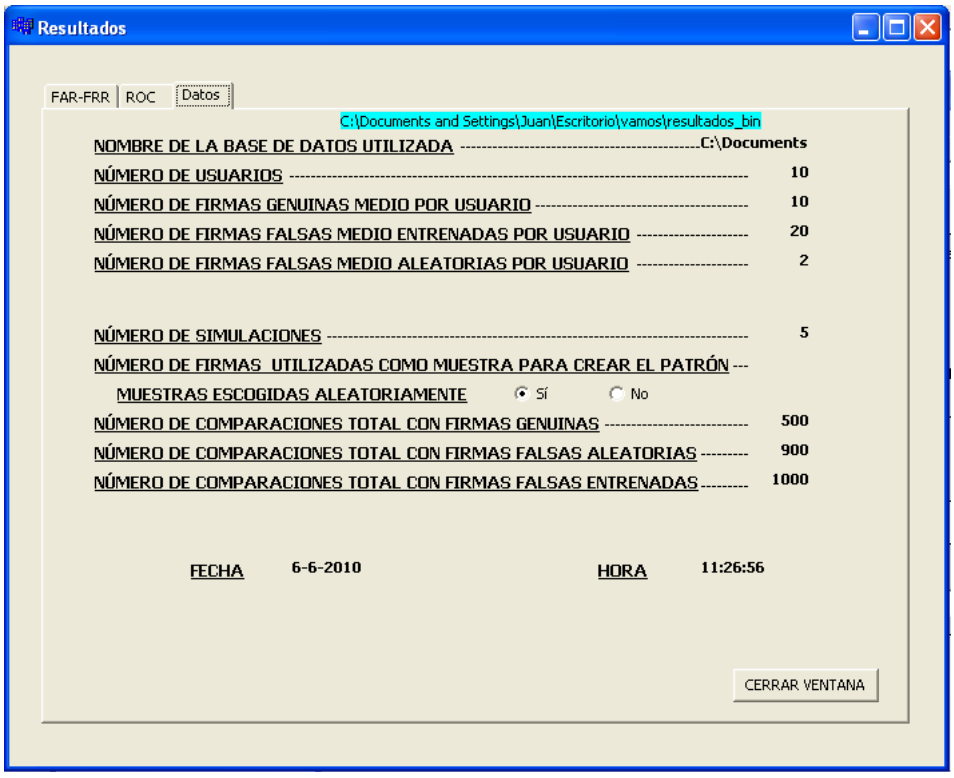


Figura 8.15.: Datos de simulación

Esta es una ventana donde vienen resumidos los datos que se introducen en la ventana principal para hacer una simulación o que estén guardados en un fichero en formato binario de simulaciones anteriores.

En los siguientes apartados de este capítulo se van a presentar algunos resultados de simulaciones realizadas con varias de las bases de datos que se han mencionado en este documento.

8.6.5.1. Resultados DTW_C_MCyT (con firmas falsas aleatorias)

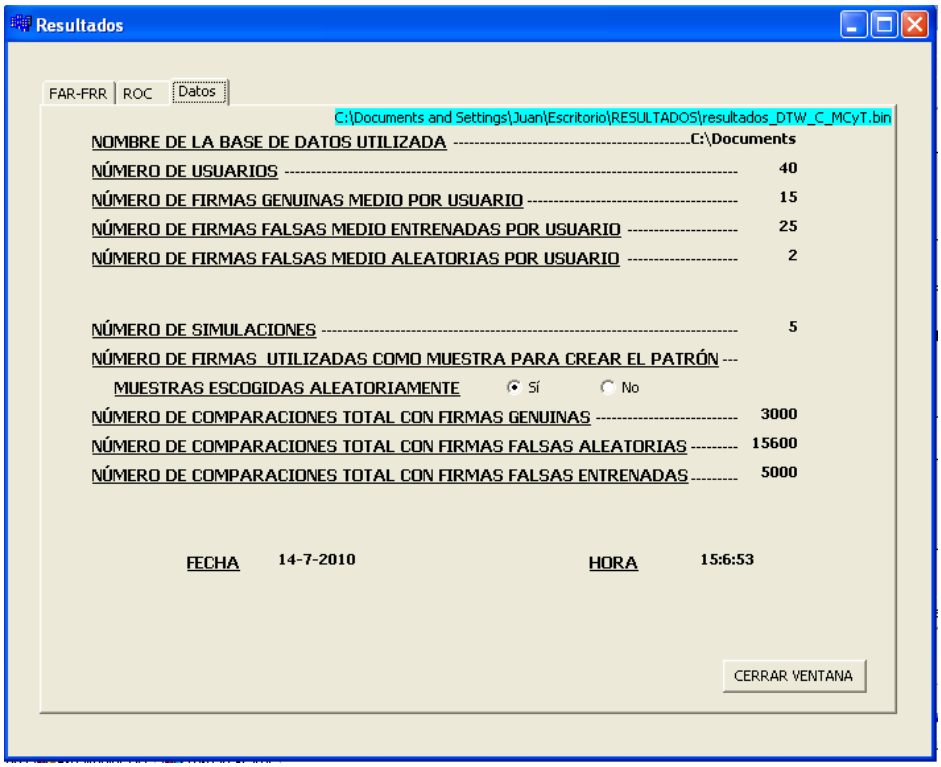


Figura 8.16.: Datos de simulación

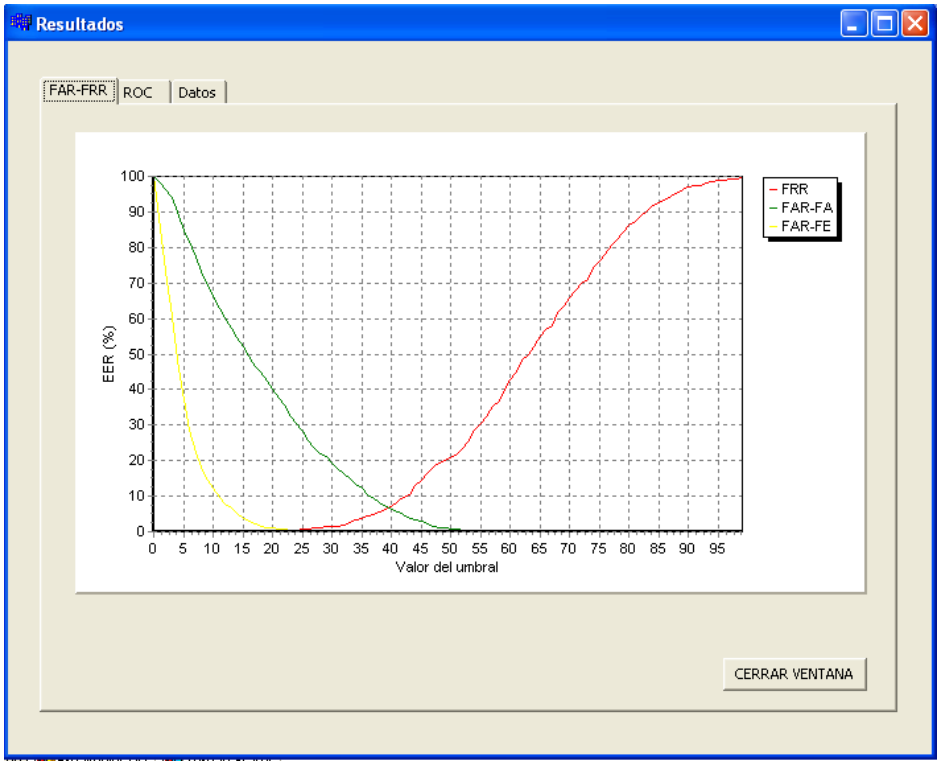


Figura 8.17.: Gráfica FAR-FRR

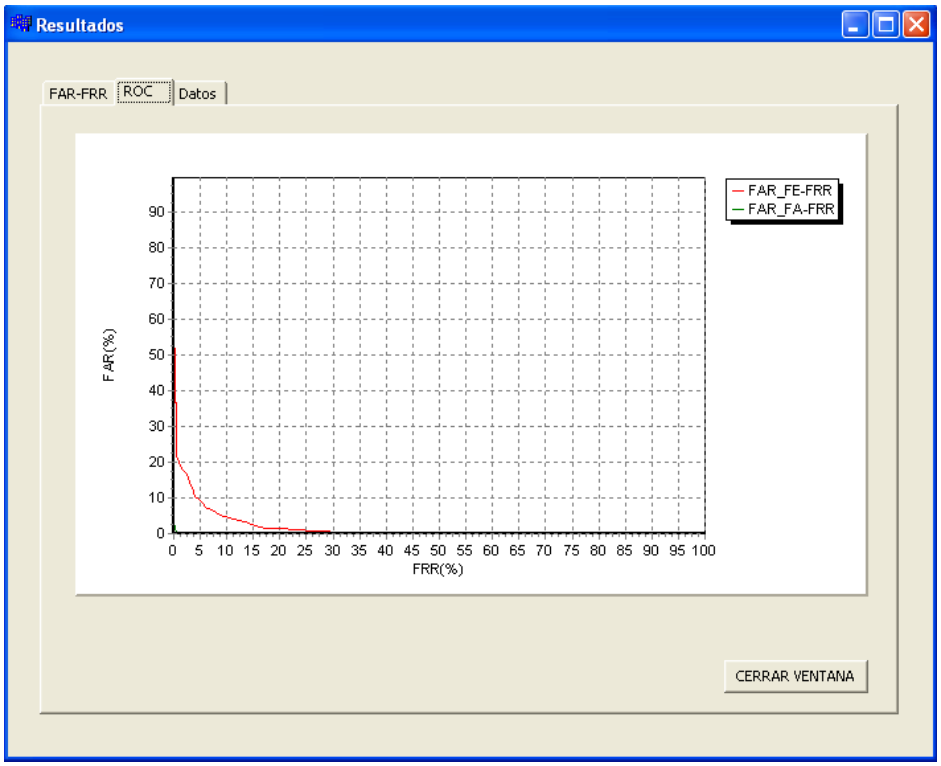


Figura 8.18.: Gráfica ROC

Este primer resultado corresponde a la base de datos MCyT simulada con el algoritmo DTW. En la ventana correspondiente a los datos de la simulación vemos los valores más importantes que se introdujeron en la aplicación para simular los resultados que se obtienen en las dos gráficas. Se observa que sí se han utilizado firmas falsas aleatorias.

8.6.5.2. Resultados DTW_C_SVC (con firmas falsas aleatorias)

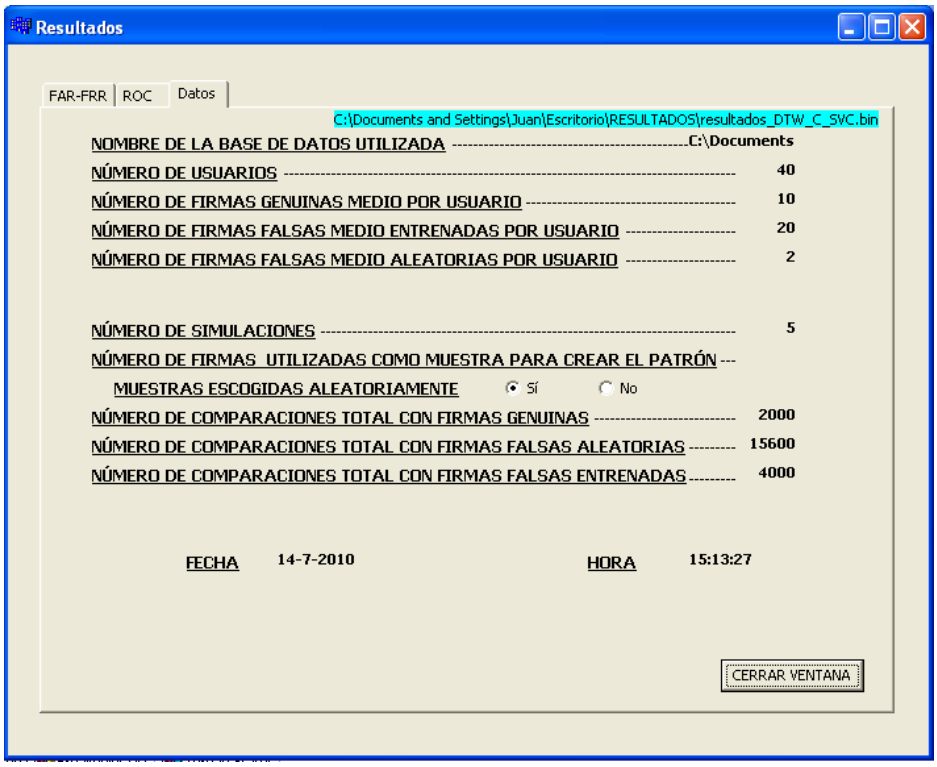


Figura 8.19.: Datos de la simulación

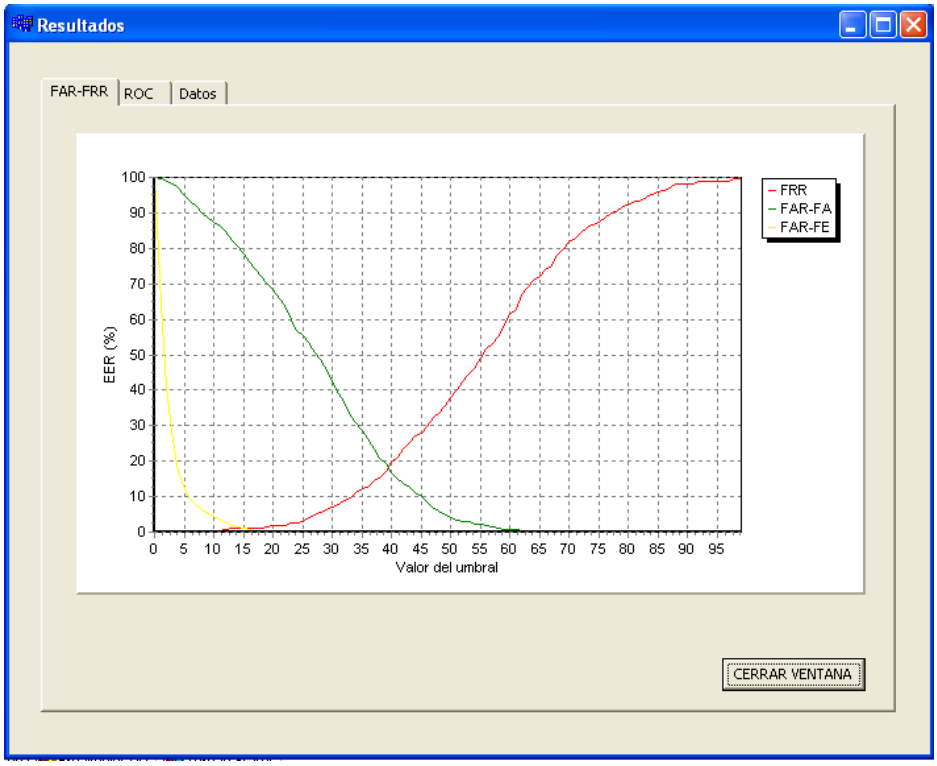


Figura 8.20.: Gráfica FAR-FRR

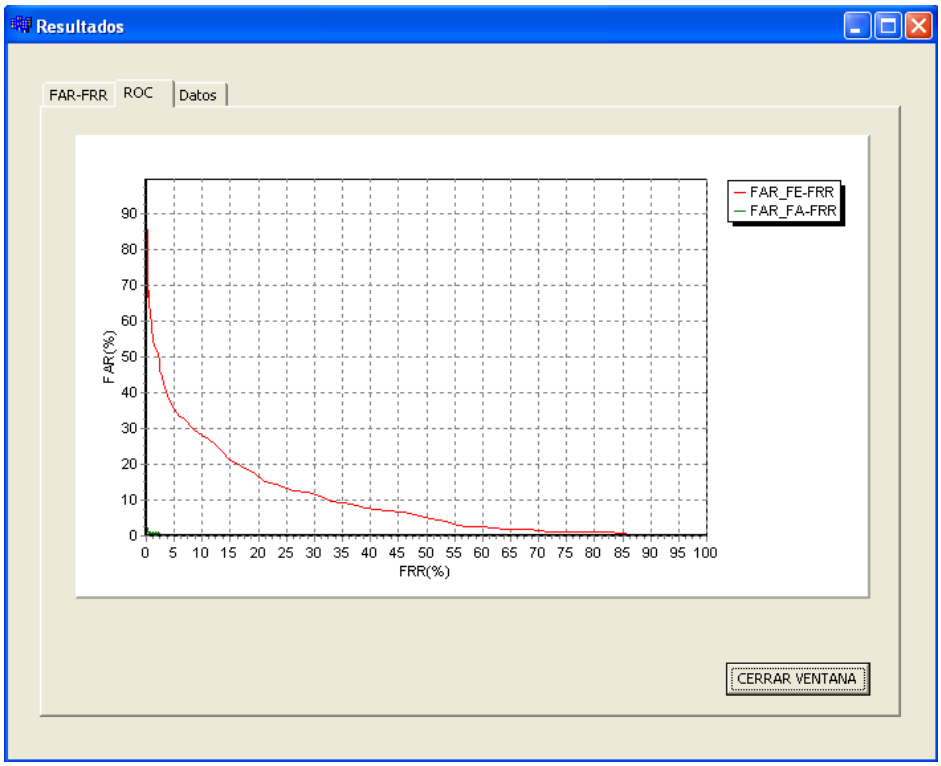


Figura 8.21.: Gráfica ROC

Este resultado corresponde a la base de datos SVC simulada con el algoritmo DTW. En la ventana correspondiente a los datos de la simulación vemos los valores más importantes que se introdujeron en la aplicación para simular los resultados que se obtienen en las dos gráficas. Se observa que sí se han utilizado firmas falsas aleatorias.

8.6.5.3. Resultados GMM_MCyT (con firmas falsas aleatorias)

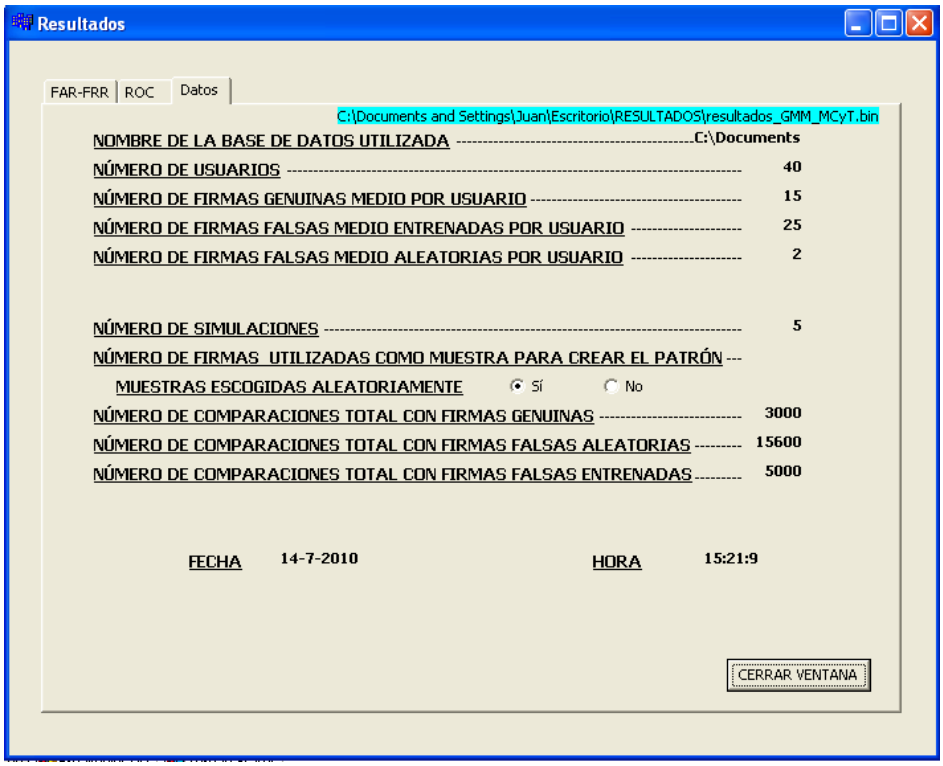


Figura 8.22.: Datos de la simulación

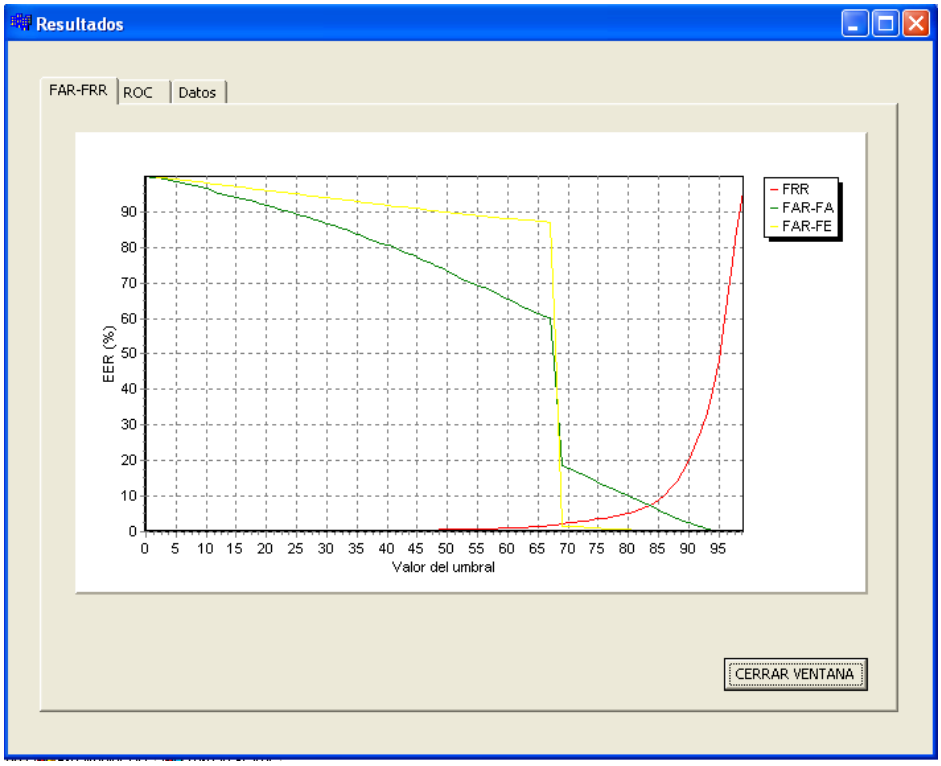


Figura 8.23.: Gráfica FAR-FRR

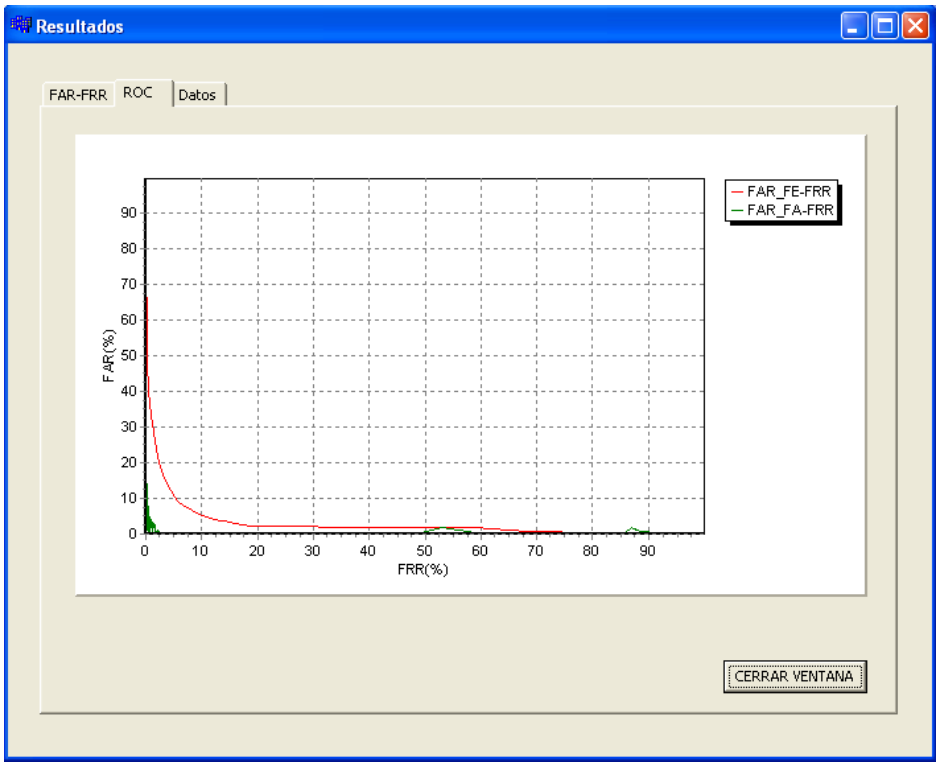


Figura 8.24.: Gráfica ROC

Este resultado corresponde a la base de datos MCyT simulada con el algoritmo GMM. En la ventana correspondiente a los datos de la simulación vemos los valores más importantes que se introdujeron en la aplicación para simular los resultados que se obtienen en las dos gráficas. Se observa que sí se han utilizado firmas falsas aleatorias.

8.6.5.4. Resultados DTW_SVC (sin firmas falsas aleatorias)

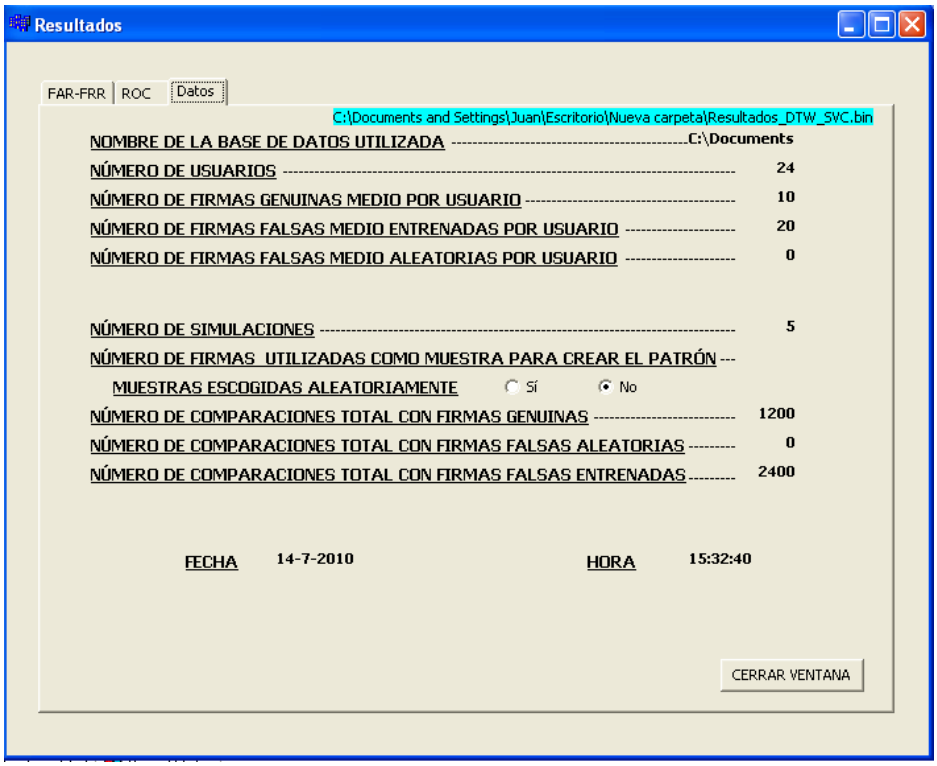


Figura 8.25.: Datos de la simulación

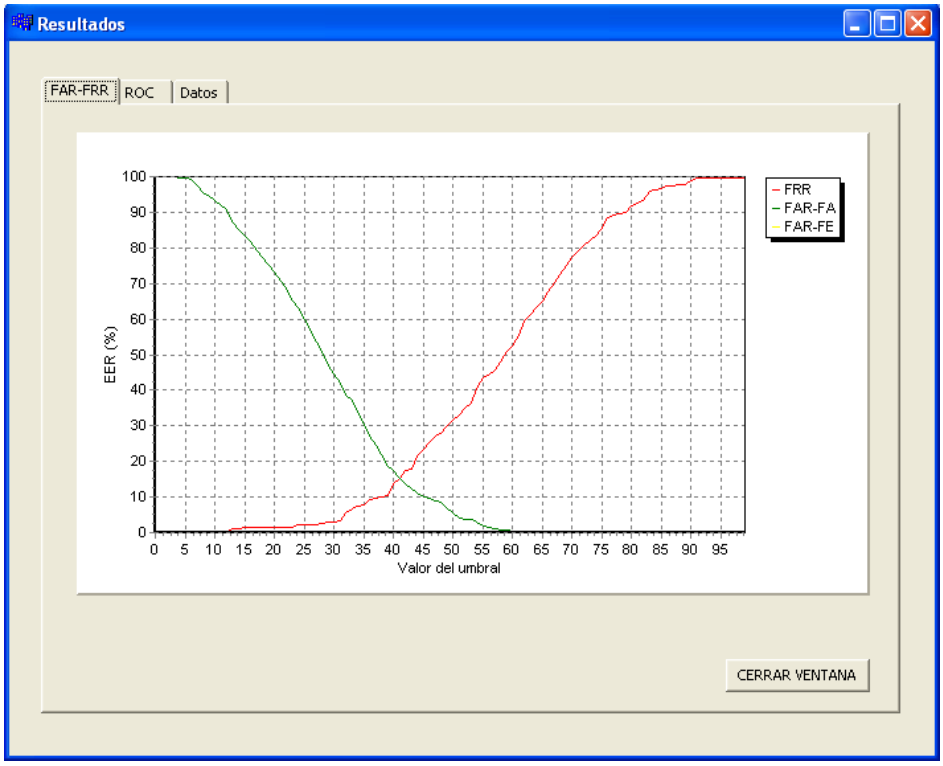


Figura 8.26.: Gráfica FAR-FRR

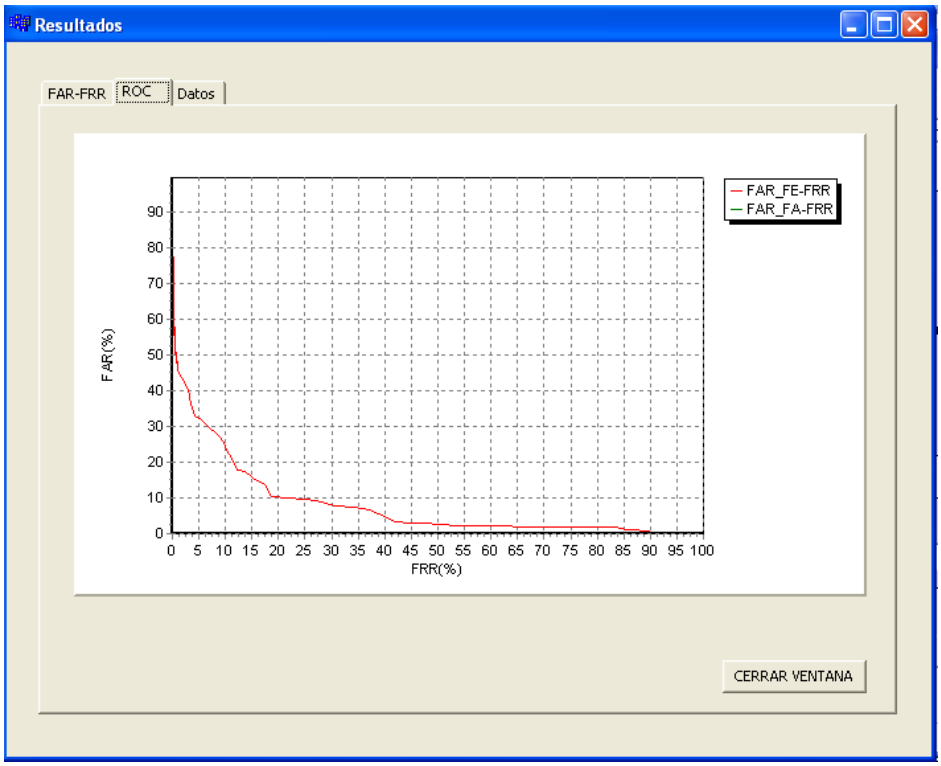


Figura 8.27.: Gráfica ROC

Este resultado corresponde a la base de datos SVC simulada con el algoritmo DTW. En la ventana correspondiente a los datos de la simulación vemos los valores más importantes que se introdujeron en la aplicación para simular los resultados que se obtienen en las dos gráficas. Se observa que no se han utilizado firmas falsas aleatorias.

9

9. CONCLUSIONES Y TRABAJOS FUTUROS

En el capítulo 1 de este documento se expusieron los objetivos de este Proyecto Fin de Carrera que como se resume en el título consistía en el diseño de una interfaz gráfica de evaluación de algoritmos de firma manuscrita.

Las funciones principales que se trataba de realizar con la aplicación eran las siguientes:

1. Introducir nuevas bases de datos de firmas de usuarios en memoria.
2. Eliminar bases de datos de firmas de usuarios.
3. Mostrar información de bases de datos de firmas de usuarios.
4. Mostrar información de los algoritmos de evaluación de firmas.
5. Seleccionar una base de datos de firmas de usuarios existente y un algoritmo para realizar una evaluación bajo una serie de parámetros.
6. Cargar resultados de simulaciones anteriores guardadas en ficheros con formato binario.
7. Guardar los resultados obtenidos de una simulación.

A la hora de presentar este Proyecto podemos decir que los objetivos se han cumplido. Esta aplicación desarrollada puede ser utilizada por investigadores para evaluar sus algoritmos o las bases de datos de firmas manuscritas recogidas. En el capítulo anterior se ha analizado más en detalle cada uno de los siete puntos que resumían las funciones de la aplicación, mostrando en cada caso los menús y ventanas que se han programado para cumplir tal propósito. El resultado se puede considerar satisfactorio a la vista de los resultados de casos concretos con algoritmos y bases de datos concretos que también se presentaron también al concluir el anterior capítulo.

Como conclusión de este trabajo tendría que decir que el campo de la Biometría y concretamente el campo de la identificación mediante firma manuscrita, que ha sido el centro del Proyecto, tiene que ser todavía muy estudiado y desarrollado. Existen ya muchos trabajos, artículos publicados y productos que se comercializan basados en esta rama de la Biometría. Algunos trabajos realizados han servido para el desarrollo de la aplicación o la elaboración de este documento.

Este Proyecto admitiría mejoras posteriores según se vaya viendo la necesidad de los investigadores del sector.

- Se podría mejorar la gestión de los algoritmos de firma manuscrita permitiendo al usuario dar de alta a un nuevo algoritmo o a dar de baja a otro que ya no le interese (se podría incluir en el menú de algoritmos ventanas del estilo a las disponibles para las bases de datos que permitieran estas funciones de alta y baja).
- Se podría dar la opción de incluir los algoritmos de firma a través de dll (librería de enlace dinámico).
- Se podría adaptar la aplicación para que los investigadores trabajen con BSP (Biometric Solution Provider), de manera que por ejemplo se pueda crear una base de datos relacionando una tableta gráfica con la aplicación.
- Se podría dar la opción al usuario de configurar los parámetros internos de la aplicación, como puede ser establecer capacidades máximas de usuarios y firmas disponibles para cada usuario.
- Se podría crear una ventana para la configuración de parámetros del algoritmo. Para esto, los algoritmos de identificación deberían enviar a la aplicación que parámetros se pueden configurar, el tipo de parámetros y su rango de valores.
- Se podría dar la opción de generar un informe en formato PDF donde se indiquen todos los datos utilizados en la simulación de una base de datos.
- Se podría dar la opción de trabajar con más de una base de datos a la vez o hacer varias simulaciones en paralelo para comparar los resultados obtenidos para un mismo algoritmo.

La identificación mediante firma manuscrita a través de Sistemas de Identificación Biométrica no está todavía muy implantada en nuestra sociedad pero se están obteniendo grandes avances.
El desarrollo de nuevos algoritmos más eficientes, la captura de nuevas características de la firma, el desarrollo de la tecnología para la captura de las firmas, el tratamiento de los datos, el desarrollo de nuevos test de evaluación, etc. Todo esto son los campos que los investigadores quieren cubrir para garantizar la seguridad del Sistema de Identificación Biométrica que debe ser el principal objetivo a cumplir.

10

10. REFERENCIAS

Para la elaboración de este documento me he servido de la gran fuente de conocimiento que supone Internet, gracias a la cual he tenido acceso a mucha de la información que he tratado de exponer en los distintos capítulos que conforman la memoria final. También he tenido acceso a trabajos y artículos publicados sobre Biometría.
A continuación quiero presentar una lista de los documentos que he podido manejar para escribir estas líneas.

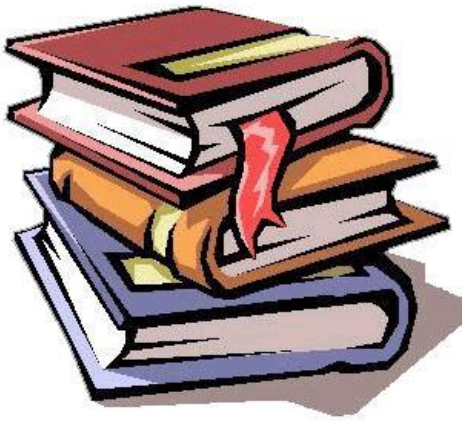
- J. Ortega-Garcia, et al., "MCYT baseline corpus: a bimodal biometric database," Vision, Image and Signal Processing, IEE Proceedings -, vol. 150, pp. 395-401, 2003.
- D. Y. Yeung, et al., "SVC2004: First international signature verification competition," Biometric Authentication, Proceedings, vol. 3072, pp. 16-22, 2004.
- B. Dumas, et al., "MyIdea - Multimodal Biometrics Database, Description of Acquisition Protocols," in Third COST 275 Workshop (COST 275), Hatfield (UK), 2005, pp. 59-62.
- O. Miguel-Hurtado, et al., "A new algorithm for signature verification system based on DTW and GMM," in Security Technology, 2008. ICCST 2008. 42nd Annual IEEE International Carnahan Conference on, 2008, pp. 206-213.
- O. Miguel-Hurtado, et al., "On-Line Signature Verification by Dynamic Time Warping and Gaussian Mixture Models," in Security Technology, 2007 41st Annual IEEE International Carnahan Conference on, 2007, pp. 23-29.

- O. Miguel-Hurtado, et al., "A new algorithm for signature verification system based on DTW and GMM," in Security Technology, 2008. ICCST 2008. 42nd Annual IEEE International Carnahan Conference on, 2008, pp. 206-213.
- O. Miguel-Hurtado, et al., "On-Line Signature Verification by Dynamic Time Warping and Gaussian Mixture Models," in Security Technology, 2007 41st Annual IEEE International Carnahan Conference on, 2007, pp. 23-29.
- J. Ortega-Garcia, et al., "MCYT baseline corpus: a bimodal biometric database," Vision, Image and Signal Processing, IEE Proceedings -, vol. 150, pp. 395-401, 2003.
- D. Y. Yeung, et al., "SVC2004: First international signature verification competition," Biometric Authentication, Proceedings, vol. 3072, pp. 16-22, 2004.
- B. Dumas, et al., "MyIdea - Multimodal Biometrics Database, Description of Acquisition Protocols," in Third COST 275 Workshop (COST 275), Hatfield (UK), 2005, pp. 59-62.

- O. Miguel-Hurtado, et al., "Analysis on compact data formats for the performance of handwritten signature biometrics," in Security Technology, 2009. 43rd Annual 2009 International Carnahan Conference on, 2009, pp. 339-346.
- Information technology — Biometric data interchange formats —Part 7:Signature/sign time series data ISO/IEC FDIS 19794-7,2006.
- A.J. Mansfield and J.L.Wayman, “Best Practices in Testing and Reporting Performance of Biometric Devices”, Aug.2002.
- Information Technology—Biometric performance testing and reporting—Part 1: Principles and Framework, ISO Standard ISO/IEC FDIS 19795-1, 2005.
- Information Technology—Biometric performance testing and reporting—Part 2: Testing Methodologies for Technology and Scenario Evaluation, ISO Standard ISO/IEC FDIS 19795-2, 2006.

También se ha utilizado información asociada a empresas y organizaciones encargadas en el desarrollo de aplicaciones para Sistemas de Identificación Biométrica:

- www.iso.org : portal de la International Organization for Standardization.
- http://www.signplus.com/en/: portal de una empresa que ofrece soluciones biométricas dentro del campo de la firma manuscrita.
- http://www.kecrypt.com/index.php: portal de una empresa que ofrece soluciones biométricas dentro del campo de la firma manuscrita.
- http://www.uc3m.es/portal/page/portal/grupos_investigacion/guti: página de la Universidad Carlos III con información referida al Grupo de Investigación G.U.T.I.



A

A.PRESUPUESTO

Para elaborar el presupuesto se han tenido en cuenta tanto los recursos materiales como los humanos. La duración del proyecto ha sido de 12 meses y la tasa de costes indirectos que se ha aplicado es del 10%.
En la plantilla adjunta en este anexo se aprecian más en detalle cada uno de los gastos involucrados en el proyecto y que en los siguientes apartados se trata de analizar brevemente.

A.1.PERSONAL

Se ha contado con un Ingeniero Técnico trabajando 3,5 horas diarias. Esto corresponde a una dedicación del 50% por eso en la plantilla del presupuesto se puede observar que se ha puesto 6 hombres mes.
Además se ha contado con los servicios de una secretaria para las labores de mecanografiado o para hacer fotocopias entre otras funciones. Esta secretaria ha realizado su trabajo en el último mes del proyecto, por eso en la tabla de personal sólo aparece 1 hombre mes.
Los costes de estos empleados aparecen en la tabla referida a personal.

A.2.EQUIPOS

Dentro de este apartado se va a tener en cuenta el material tanto de hardware como de software utilizado.
El equipo sobre el cual se ha trabajado ha sido un ordenador portátil de la marca HP modelo Pavilion dv6000. El precio de este ordenador fue de 600€ y al dedicarlo durante 12 meses al proyecto se estima un valor de amortización de 120€.
Se ha comprado una licencia de Borland C++ para el desarrollo de la aplicación con un precio de 1281,62€ cuyo valor de amortización en los 12 meses de proyecto queda en 256,32€.

El ordenador portátil venía con una licencia de Microsoft Office 2007 que ha servido para la elaboración de informes, memorias, presentaciones powerpoint, entre otros trabajos de oficina. El coste de esta licencia por tanto es nulo.
También se ha tenido que acceder a las bases de datos de firmas manuscrita, también con coste nulo al ser utilizadas para investigación dentro de la Universidad.

El total de los costes de equipos para el periodo de realización del proyecto asciende a 376,32€.

A.3.SUBCONTRATACIÓN DE TAREAS

Para la impresión y encuadernación de la memoria del proyecto se han solicitado los servicios de la empresa CopyServi. Los costes de estas tareas ascienden a los 200€.

No se ha necesitado ningún otro trabajo de otras empresas.

A.4.OTROS COSTES DIRECTOS DEL PROYECTO

En este apartado se han contado los gastos fungibles como son los materiales de oficina. El gasto asciende a los 30€.

A.5.COSTES INDIRECTOS

Se aplica una tasa del 10% para el cálculo de los costes indirectos que se hace sobre el total de los costes directos. En estos costes indirectos se van a incluir los gastos de electricidad e internet necesarios para elaborar el proyecto.

El total de los costes indirectos asciende a los 1742€.

A.6.PRESUPUESTO DE PROYECTO

UNIVERSIDAD CARLOS III DE MADRID
Escuela Politécnica Superior

PRESUPUESTO DE PROYECTO

1.- Autor: Juan Fernández García-Obledo

2.- Departamento: Tecnología Electrónica

3.- Descripción del Proyecto:

- Titulo: **Diseño de una interfaz de evaluación de algoritmos de firma manuscrita**
- Duración (meses): **12**
- Tasa de costes Indirectos: 10%

4.- Presupuesto total del Proyecto (valores en Euros):

Euros

5.- Desglose presupuestario (costes directos)

PERSONAL

Apellidos y nombre	N.I.F. (no rellenar - solo a titulo informativo)	Categoría	Dedicación [a] (hombres mes)	Coste hombre mes	Coste (Euro)	Firma de conformidad
					0,00	
		Ingeniero Senior		4.289,54	0,00	
Juan Fernández García-Obledo		Ingeniero	6	2.694,39	16.166,34	
Marta Sanz Ruiz		Secretaria	1	644,76	644,76	
					0,00	
		Hombres mes	**7**	**Total**	16.811,10	

[a] 1 Hombre mes = 131,25 horas. Máximo anual de dedicación de 12 hombres mes (1575 horas)
Máximo anual para PDI de la Universidad Carlos III de Madrid de 8,8 hombres mes (1.155 horas)

EQUIPOS

Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable [d]
Ordenador Portátil HP Pavilion dv600	600,00	100	12	60	120,00
C++ Builder	1.281,62	100	12	60	256,32
Microsoft Office 2007	0,00	100	12	60	0,00
Bases de Datos Firma Manuscrita	0,00	100	12	60	0,00
		100		60	0,00
					0,00
				Total	376,32

[d] Fórmula de cálculo de la Amortización:

$$\frac{A}{B} x C x D$$

A = nº de meses desde la fecha de facturación en que el equipo es utilizado
B = periodo de depreciación (60 meses)
C = coste del equipo (sin IVA)
D = % del uso que se dedica al proyecto (habitualmente 100%)

SUBCONTRATACIÓN DE TAREAS

Descripción	Empresa	Coste imputable
Impresión y encuadernación memori	CopyServi	200,00
	Total	200,00

OTROS COSTES DIRECTOS DEL PROYECTO[e]

Descripción	Empresa	Costes imputable
Material de oficina		30,00
	Total	30,00

[e] Este capítulo de gastos incluye todos los gastos no contemplados en los conceptos anteriores, por ejemplo: fungible, viajes y dietas, otros,...

6.- Resumen de costes

Presupuesto Costes Totales	Presupuesto Costes Totales
Personal	16.811
Amortización	376
Subcontratación de tareas	200
Costes de funcionamiento	30
Costes Indirectos	1.742
Total	19.159